VISIT…

Jenő Sólyom

Fundamentals of the Physics of Solids

Volume 1 – Structure and Dynamics

Springer

Periodic table of elements

	1 IA												13 IIIA	14 IVA	15 VA	16 VIA	17 VIIA	18 VIIIA
1	1 H $1s^1$ Hydrogen	2 IIA																2 He $1s^2$ Helium
2	3 Li [He] $2s^1$ Lithium	4 Be $1s^2$ Beryllium											5 B $2s^2 2p^1$ Boron	6 C $2s^2 2p^2$ Carbon	7 N $2s^2 2p^3$ Nitrogen	8 O $2s^2 2p^4$ Oxygen	9 F $2s^2 2p^5$ Fluorine	10 Ne $2s^2 2p^6$ Neon
3	11 Na [Ne] $3s^1$ Sodium	12 Mg $3s^2$ Magnesium	3 IIIB	4 IVB	5 VB	6 VIB	7 VIIB	8	9 VIIIB	10	11 IB	12 IIB	13 Al $3s^2 3p^1$ Aluminum	14 Si $3s^2 3p^2$ Silicon	15 P $3s^2 3p^3$ Phosphorus	16 S $3s^2 3p^4$ Sulfur	17 Cl $3s^2 3p^5$ Chlorine	18 Ar $3s^2 3p^6$ Argon
4	19 K [Ar] $4s^1$ Potassium	20 Ca $4s^2$ Calcium	21 Sc $3d^1 4s^2$ Scandium	22 Ti $3d^2 4s^2$ Titanium	23 V $3d^3 4s^2$ Vanadium	24 Cr $3d^5 4s^1$ Chromium	25 Mn $3d^5 4s^2$ Manganese	26 Fe $3d^6 4s^2$ Iron	27 Co $3d^7 4s^2$ Cobalt	28 Ni $3d^8 4s^2$ Nickel	29 Cu $3d^{10} 4s^1$ Copper	30 Zn $3d^{10} 4s^2$ Zinc	31 Ga $3d^{10} 4s^2 4p^1$ Gallium	32 Ge $3d^{10} 4s^2 4p^2$ Germanium	33 As $3d^{10} 4s^2 4p^3$ Arsenic	34 Se $3d^{10} 4s^2 4p^4$ Selenium	35 Br $3d^{10} 4s^2 4p^5$ Bromine	36 Kr $3d^{10} 4s^2 4p^6$ Krypton
5	37 Rb [Kr] $5s^1$ Rubidium	38 Sr $5s^2$ Strontium	39 Y $4d^1 5s^2$ Yttrium	40 Zr $4d^2 5s^2$ Zirconium	41 Nb $4d^4 5s^1$ Niobium	42 Mo $4d^5 4s^1$ Molybdenum	43 Tc $4d^5 5s^2$ Technetium	44 Ru $4d^7 5s^1$ Ruthenium	45 Rh $4d^8 5s^1$ Rhodium	46 Pd $4d^{10}$ Palladium	47 Ag $4d^{10} 5s^1$ Silver	48 Cd $4d^{10} 5s^2$ Cadmium	49 In $4d^{10} 5s^2 5p^1$ Indium	50 Sn $4d^{10} 5s^2 5p^2$ Tin	51 Sb $4d^{10} 5s^2 5p^3$ Antimony	52 Te $4d^{10} 5s^2 5p^4$ Tellurium	53 I $4d^{10} 5s^2 5p^5$ Iodine	54 Xe $4d^{10} 5s^2 5p^6$ Xenon
6	55 Cs [Xe] $6s^1$ Cesium	56 Ba $6s^2$ Barium	57 La $5d^1 6s^2$ Lanthanum	72 Hf $5d^2 6s^2$ Hafnium	73 Ta $5d^3 6s^2$ Tantalum	74 W $5d^4 6s^2$ Tungsten	75 Re $5d^5 6s^2$ Rhenium	76 Os $5d^6 6s^2$ Osmium	77 Ir $5d^7 6s^2$ Iridium	78 Pt $5d^9 6s^1$ Platinum	79 Au $5d^{10} 6s^1$ Gold	80 Hg $5d^{10} 6s^2$ Mercury	81 Tl $5d^{10} 6s^2 6p^1$ Thallium	82 Pb $5d^{10} 6s^2 6p^2$ Lead	83 Bi $5d^{10} 6s^2 6p^3$ Bismuth	84 Po $5d^{10} 6s^2 6p^4$ Polonium	85 At $5d^{10} 6s^2 6p^5$ Astatine	86 Rn $5d^{10} 6s^2 6p^6$ Radon
7	87 Fr [Rn] $7s^1$ Francium	88 Ra $7s^2$ Radium	89 Ac $6d^1 7s^2$ Actinium	104 Rf $6d^2 7s^2$ Rutherfordium	105 Db $6d^3 7s^2$ Dubnium	106 Sg $6d^4 7s^2$ Seaborgium	107 Bh $6d^5 7s^2$ Bohrium	108 Hs $6d^6 7s^2$ Hassium	109 Mt $6d^7 7s^2$ Meitnerium	110 Ds $6d^9 7s^1$ Darmstadtium	111 Rg $6d^{10} 7s^1$ Roentgenium							

Lanthanoids	58 Ce $4f^2 6s^2$ Cerium	59 Pr $4f^3 6s^2$ Praseodymium	60 Nd $4f^4 6s^2$ Neodymium	61 Pm $4f^5 6s^2$ Promethium	62 Sm $4f^6 6s^2$ Samarium	63 Eu $4f^7 6s^2$ Europium	64 Gd $4f^7 5d^1 6s^2$ Gadolinium	65 Tb $4f^9 6s^2$ Terbium	66 Dy $4f^{10} 6s^2$ Dysprosium	67 Ho $4f^{11} 6s^2$ Holmium	68 Er $4f^{12} 6s^2$ Erbium	69 Tm $4f^{13} 6s^2$ Thulium	70 Yb $4f^{14} 6s^2$ Ytterbium	71 Lu $4f^{14} 5d^1 6s^2$ Lutetium
Actinoids	90 Th $6d^2 7s^2$ Thorium	91 Pa $5f^2 6d^1 7s^2$ Protactinium	92 U $5f^3 6d^1 7s^2$ Uranium	93 Np $5f^5 7s^2$ Neptunium	94 Pu $5f^6 7s^2$ Plutonium	95 Am $5f^7 7s^2$ Americium	96 Cm $5f^7 6d^1 7s^2$ Curium	97 Bk $5f^8 6d^1 7s^2$ Berkelium	98 Cf $5f^{10} 7s^2$ Californium	99 Es $5f^{11} 7s^2$ Einsteinium	100 Fm $5f^{12} 7s^2$ Fermium	101 Md $5f^{13} 7s^2$ Mendelevium	102 No $5f^{14} 7s^2$ Nobelium	103 Lr $5f^{14} 6d^1 7s^2$ Lawrencium

Fundamentals of the Physics of Solids

Jenő Sólyom

Fundamentals of the Physics of Solids

Volume I

Structure and Dynamics

Translated by Attila Piróth

With 240 Figures and 50 Tables

Springer

Author
Professor Jenő Sólyom

Research Institute for Solid State Physics
and Optics
P.O.Box 49
1525 Budapest
Hungary
and
Department of Physics
Eötvös Loránd University
1117 Budapest
Pázmány sétány 1/A
Hungary
solyom@szfki.hu

Translator
Attila Piróth
www.pirothattila.com

Title of the Hungarian edition: *A Modern Szilárdtestfizika Alapjai I. A szilárd testek szerkezete és dinamikája*, ISBN 963 463 538 5, ©ELTE Eötvös Kiadó 2002

Library of Congress Control Number: 2007929725

ISBN 978-3-540-72599-2 Springer Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable for prosecution under the German Copyright Law.

Springer is a part of Springer Science+Business Media
springer.com
© Springer-Verlag Berlin Heidelberg 2007

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Typesetting: by the author, the translator, and Integra, India using a Springer LaTeX macro package
Cover design: F. Steinen, eStudio Calamar, Girona/Spain.

Printed on acid-free paper SPIN: 11429470 5 4 3 2 1 0

To Márta, Gyöngyvér, Tünde, and lringó

Preface

The reader is holding the first volume of a three-volume textbook on solid-state physics. This book is the outgrowth of the courses I have taught for many years at Eötvös University, Budapest, for undergraduate and graduate students under the titles *Solid-State Physics* and *Modern Solid-State Physics*.

The main motivation for the publication of my lecture notes as a book was that none of the truly numerous textbooks covered all those areas that I felt should be included in a multi-semester course. Especially, if the course strives to present solid-state physics in a unified structure, and aims at discussing not only classic chapters of the subject matter but also (in more or less detail) problems that are of great interest for today's researcher as well. Besides, the book presents a much larger material than what can be covered in a two- or three-semester course. In the first part of the first volume the analysis of crystal symmetries and structure goes into details that certainly cannot be included in a usual course on solid-state physics. The same applies, among others, to the discussion of the methods used in the determination of band structure, the properties of Fermi liquids and non-Fermi liquids, and the theory of unconventional superconductors in the second and third volumes. These parts can be assigned as supplementary reading for interested students, or can be discussed in advanced courses.

The line of development and the order of the chapters are based on the prerequisites for understanding each part. Therefore a gradual shift can be observed in the style of the book. While the intermediate steps of calculations are presented in considerable detail and explanations are also more lengthy in the first and second volumes, they are much sparser and more concise in the third one, thus this volume relies more on the individual work of the students. On account of the prerequisites, certain topics have to be revisited. This is why magnetic properties are treated in three, and superconductivity in two parts. The magnetism of individual atoms is presented in an introductory chapter. The structure and dynamics of magnetically ordered systems built up of localized moments are best discussed after lattice vibrations, along the same lines. Magnetism is revisited in the third volume, where the role of electron–electron

interactions is discussed in more detail. Similarly, the phenomenological description of superconductivity is presented after the analysis of the transport properties of normal metals, in contrast to them, while the microscopic theory is outlined later, when the effects of interactions are discussed.

Separating the material into three similar-sized volumes is a necessity in view of the size of the material – but it also reflects the internal logical structure of the subject matter. At those universities where the basic course in solid-state physics runs for three semesters working through one volume per semester is a natural schedule. In this case the discussion of the electron gas – which is traditionally part of the introduction – is left for the second semester. This choice is particularly suited to curricula in which the course on solid-state physics is held parallel with quantum mechanics or statistical physics. If the lecturer feels more comfortable with the traditional approach, the discussion of the Drude model presented in the second volume can be moved to the beginning of the whole course. Nevertheless the discussion of the Sommerfeld model should be postponed until students have familiarized themselves with the fundamentals of statistical physics. For the same reason the lecturer may prefer to change the order of other chapters as well. Apart from the presentation of the consequences of translational symmetry, the topics discussed in Chapters 3, 4, and 6 can be deferred to a later time, when students have acquired a sound knowledge of quantum mechanics, atomic, and molecular physics. The consequences of translational symmetry can also be analyzed after the discussion of phonons. All this is, to a large extent, up to the personal preferences of the lecturer.

In presenting the field of solid-state physics, special emphasis has been laid on discussing the physical phenomena that can be observed in solids. Nevertheless I have tried to give – or at least outline – the theoretical interpretation for each phenomenon, too. As is common practice for textbooks, I have omitted precise references that would give the publication data of the discussed results. I have made exceptions only for figures taken directly from published articles. At the end of the first, introductory chapter I have given a list of textbooks and series on solid-state physics, while at the end of each subsequent chapter I have listed textbooks and review articles that present further details and references pertaining to the subject matter of the chapter in question.

Bulky as it might be, this three-volume treatise presents only the fundamentals of solid-state physics. Today, when articles about condensed matter physics fill tens of thousands of pages every year in Physical Review alone, it would be obviously overambitious to aim at more. Therefore, building on the foundations presented in this book students will have to acquire a substantial amount of extra knowledge before they can understand the subtleties of the topics in the forefront of today's research. Nevertheless at the end of the third volume students will also appreciate the number of open questions and the necessity of further research.

A certain knowledge of quantum mechanics is a prerequisite for studying solid-state physics. Various techniques of quantum mechanics – above all field-theoretical methods and methods employed in solving many-body problems – play an important role in present-day solid-state physics. Some essential details are listed in one of the appendices of the third volume, however, I have omitted more complicated calculations that would have required the application of the modern apparatus of many-body problems. This is especially true for the third volume, where central research topics of present-day solid-state physics are discussed, in which the theoretical interpretation of experimental results is often impossible without some extremely complex mathematical formulation.

The selection of topics obviously bears the stamp of the author's own research interest, too. This explains why the discussion of certain important fields – such as the mechanical properties of solids, surface phenomena, amorphous systems or mesoscopic systems, to name but a few – have been omitted.

I have used the International System of Units (SI), and have given the equations of electromagnetism in rationalized form. Nonrationalized equations as well as gaussian CGS (and other) units are nevertheless still very much in use in the solid-state physics literature. This has been indicated at the appropriate places. On a few occasions I have also given the formulas obtained from nonrationalized equations. In addition to the fundamental physical constants used in solid-state physics, the commonest conversion factors are also listed in Appendix A. Only once have I deviated from standard practice, denoting Boltzmann's constant by k_{B} instead of k – reserving the latter for the wave number, which plays a central role in solid-state physics.

To give an impression of the usual values of the quantities occurring in solid-state physics, typical calculated values or measured data are often tabulated. To provide the most precise data available, I have relied on the Landolt–Börnstein series, the CRC Handbook of Chemistry and Physics, and other renowned sources. Since these data are for information only, I have not indicated either their error or in many cases the measurement temperature, and I have not mentioned when different measurement methods lead to slightly disparate results. As a rule of thumb, the error is usually smaller than or on the order of the last digit.

I would like to thank all my colleagues who read certain chapters and improved the text through their suggestions and criticism. Particular thanks go to professors György Mihály and Attila Virosztek for reading the whole manuscript. In spite of all efforts, some mistakes have certainly remained in the book. Obviously, the author alone bears the responsibility for them.

Special thanks are due to Károly Härtlein for his careful work in drawing the majority of the figures. The figures presenting experimental results are reproduced with the permission of the authors or the publishers. M. C. Escher's drawings in Chapter 5 are reproduced with the permission of the copyright holder © 2006 The M. C. Escher Company-Holland. The challenge

of translating the book from the Hungarian original was taken up by Attila Piróth. I acknowledge his work.

Finally, I am indebted to my family, to my wife and children, for their patience during all those years when I spent evenings and weekends with writing this book.

Budapest, May 2007 Jenő Sólyom

Contents

1 Introduction 1

2 The Structure of Condensed Matter 13
2.1 Characterization of the Structure 14
2.1.1 Short- and Long-Range Order 14
2.1.2 Order in the Center-of-Mass Positions, Orientation, and Chemical Composition 19
2.2 Classification of Condensed Matter According to Structure 20
2.2.1 Solid Phase 20
2.2.2 Liquid Phase 22
2.2.3 Mesomorphic Phases 23

3 The Building Blocks of Solids 31
3.1 Solids as Many-Particle Systems 31
3.1.1 The Hamiltonian of Many-Particle Systems 32
3.1.2 Effects of Applied Fields 34
3.1.3 Relativistic Effects 36
3.2 The State of Ion Cores 38
3.2.1 Hund's Rules 41
3.2.2 Angular Momentum and Magnetic Moment 44
3.2.3 The Magnetic Hamiltonian of Atomic Electrons 46
3.2.4 Magnetization and Susceptibility 47
3.2.5 Langevin or Larmor Diamagnetism 49
3.2.6 Atomic Paramagnetism 51
3.2.7 Van Vleck Paramagnetism 60
3.2.8 Electron Spin Resonance 61
3.3 The Role of Nuclei 68
3.3.1 Interaction with Nuclear Magnetic Moments 68
3.3.2 Nuclear Magnetic Resonance 71
3.3.3 The Mössbauer effect 72

4 Bonding in Solids ... 75
4.1 Types of Bonds and Cohesive Energy ... 75
4.1.1 Classification of Solids According to the Type of the Bond ... 76
4.1.2 Cohesive Energy ... 76
4.2 Molecular crystals ... 78
4.2.1 Van der Waals Bonds in Quantum Mechanics ... 79
4.2.2 Cohesive Energy of Molecular Crystals ... 81
4.3 Ionic Bond ... 83
4.4 Covalent Bond ... 89
4.4.1 The Valence-Bond Method ... 90
4.4.2 Polar Covalent Bond ... 94
4.4.3 The Molecular-Orbital Method ... 96
4.4.4 The LCAO Method ... 97
4.4.5 Molecular Orbitals Between Different Atoms ... 100
4.4.6 Slater Determinant Form of the Wavefunction ... 101
4.4.7 Hybridized Orbitals ... 103
4.4.8 Covalent Bonds in Solids ... 105
4.5 Metallic Bond ... 106
4.6 The Hydrogen Bond ... 106

5 Symmetries of Crystals ... 109
5.1 Translational Symmetry in Crystals ... 110
5.1.1 Translational Symmetry in Finite Crystals ... 110
5.1.2 The Choice of Primitive Vectors ... 111
5.1.3 Bravais Lattice and Basis ... 113
5.1.4 Primitive Cells, Wigner–Seitz Cells, and Bravais Cells ... 114
5.1.5 Crystallographic Positions, Directions, and Planes ... 118
5.2 The Reciprocal Lattice ... 120
5.2.1 Definition of the Reciprocal Lattice ... 120
5.2.2 Properties of the Reciprocal Lattice ... 122
5.3 Rotations and Reflections ... 124
5.3.1 Symmetry Operations and Symmetry Elements ... 124
5.3.2 Point Groups ... 127
5.4 Rotation and Reflection Symmetries in Crystals ... 135
5.4.1 Rotation Symmetries of Bravais Lattices ... 135
5.4.2 Crystallographic Point Groups ... 137
5.4.3 Crystal Systems and Bravais Groups ... 138
5.4.4 Two-Dimensional Bravais-Lattice Types ... 142
5.4.5 Three-Dimensional Bravais-Lattice Types ... 146
5.4.6 The Hierarchy of Crystal Systems ... 154
5.5 Full Symmetry of Crystals ... 157
5.5.1 Screw Axes and Glide Planes ... 157
5.5.2 Point Groups of Crystals and Crystal Classes ... 160

5.5.3 Space Groups . . . 162
5.5.4 Symmetries of Magnetic Crystals . . . 166

6 Consequences of Symmetries . . . 171
6.1 Quantum Mechanical Eigenvalues and Symmetries . . . 172
6.1.1 Wigner's Theorem . . . 172
6.1.2 Splitting of Atomic Levels in Crystals . . . 173
6.1.3 Spin Contributions to Splitting . . . 179
6.1.4 Kramers' Theorem . . . 182
6.1.5 Selection Rules . . . 184
6.2 Consequences of Translational Symmetry . . . 184
6.2.1 The Born–von Kármán Boundary Condition . . . 185
6.2.2 Bloch's Theorem . . . 186
6.2.3 Equivalent Wave Vectors . . . 189
6.2.4 Conservation of Crystal Momentum . . . 191
6.2.5 Symmetry Properties of Energy Eigenstates . . . 194
6.3 Symmetry Breaking and Its Consequences . . . 199
6.3.1 Symmetry Breaking in Phase Transitions . . . 199
6.3.2 Goldstone's Theorem . . . 200

7 The Structure of Crystals . . . 203
7.1 Types of Crystal Structures . . . 203
7.2 Cubic Crystal Structures . . . 205
7.2.1 Simple Cubic Structures . . . 205
7.2.2 Body-Centered Cubic Structures . . . 210
7.2.3 Face-Centered Cubic Structures . . . 214
7.2.4 Diamond and Sphalerite Structures . . . 221
7.3 Hexagonal Crystal Structures . . . 224
7.4 Typical Sizes of Primitive Cells . . . 229
7.5 Layered and Chain-Like Structures . . . 229
7.6 Relationship Between Structure and Bonding . . . 233
7.6.1 The Structure of Covalently Bonded Solids . . . 233
7.6.2 Structures with Nondirectional Bonds . . . 235

8 Methods of Structure Determination . . . 241
8.1 The Theory of Diffraction . . . 242
8.1.1 The Bragg and Laue Conditions of Diffraction . . . 242
8.1.2 Structure Amplitude and Atomic Form Factor . . . 246
8.1.3 Diffraction Cross Section . . . 249
8.1.4 The Shape and Intensity of Diffraction Peaks . . . 252
8.1.5 Cancellation in Structures with a Polyatomic Basis . . . 255
8.1.6 The Dynamical Theory of Diffraction . . . 258
8.2 Experimental Study of Diffraction . . . 261
8.2.1 Characteristic Properties of Different Types of Radiation . . . 261

8.2.2 The Ewald Construction . . . 264
8.2.3 Diffraction Methods . . . 265
8.3 Other Methods of Structure Determination . . . 269

9 The Structure of Real Crystals . . . 273
9.1 Point Defects . . . 275
9.1.1 Vacancies . . . 275
9.1.2 Interstitials . . . 278
9.1.3 Pairs of Point Defects . . . 280
9.2 Line Defects, Dislocations . . . 283
9.2.1 Edge and Screw Dislocations . . . 284
9.2.2 The Burgers Vector . . . 286
9.2.3 Dislocations as Topological Defects . . . 288
9.2.4 Disclinations . . . 290
9.2.5 Dislocations in Hexagonal Lattices . . . 292
9.3 Planar Defects . . . 293
9.3.1 Stacking Faults . . . 293
9.3.2 Partial Dislocations . . . 294
9.3.3 Low-Angle Grain Boundaries . . . 298
9.3.4 Coincident-Site-Lattice and Twin Boundaries . . . 299
9.3.5 Antiphase Boundaries . . . 300
9.4 Volume Defects . . . 302

10 The Structure of Noncrystalline Solids . . . 303
10.1 The Structure of Amorphous Materials . . . 303
10.1.1 Models of Topological Disorder . . . 303
10.1.2 Analysis of the Short-Range Order . . . 305
10.2 Quasiperiodic Structures . . . 309
10.2.1 Periodic and Quasiperiodic Functions . . . 310
10.2.2 Incommensurate Structures . . . 312
10.2.3 Experimental Observation of Quasicrystals . . . 314
10.2.4 The Fibonacci Chain . . . 317
10.2.5 Penrose Tiling of the Plane . . . 323
10.2.6 Three-Dimensional Quasicrystals . . . 327

11 Dynamics of Crystal Lattices . . . 331
11.1 The Harmonic Approximation . . . 331
11.1.1 Second-Order Expansion of the Potential . . . 332
11.1.2 Expansion of the Energy for Pair Potentials . . . 334
11.1.3 Equations Governing Lattice Vibrations . . . 335
11.2 Vibrational Spectra of Simple Lattices . . . 337
11.2.1 Vibrations of a Monatomic Linear Chain . . . 337
11.2.2 Vibrations of a Diatomic Chain . . . 341
11.2.3 Vibrations of a Dimerized Chain . . . 345
11.2.4 Vibrations of a Simple Cubic Lattice . . . 349

11.3 The General Description of Lattice Vibrations 354
11.3.1 The Dynamical Matrix and its Eigenvalues 355
11.3.2 Normal Coordinates and Normal Modes 357
11.3.3 Acoustic and Optical Vibrations 360
11.4 Lattice Vibrations in the Long-Wavelength Limit 363
11.4.1 Acoustic Vibrations as Elastic Waves 363
11.4.2 Elastic Constants of Crystalline Materials 367
11.4.3 Elastic Waves in Cubic Crystals 371
11.4.4 Optical Vibrations in Ionic Crystals 373
11.5 Localized Lattice Vibrations 377
11.5.1 Vibrations in a Chain with an Impurity 377
11.5.2 Impurities in a Three-Dimensional Lattice 381
11.6 The Specific Heat of Classical Lattices 383

12 The Quantum Theory of Lattice Vibrations 387
12.1 Quantization of Lattice Vibrations 387
12.1.1 The Einstein Model 387
12.1.2 The Debye Model 389
12.1.3 Quantization of the Hamiltonian 390
12.1.4 The Quantum Mechanics of Harmonic Oscillators 392
12.1.5 Creation and Annihilation Operators of Vibrational Modes 394
12.1.6 Phonons as Elementary Excitations 395
12.1.7 Acoustic Phonons as Goldstone Bosons 397
12.1.8 Symmetries of the Vibrational Spectrum 397
12.2 Density of Phonon States 398
12.2.1 Definition of the Density of States 399
12.2.2 The Density of States in One- and Two-Dimensional Systems 402
12.2.3 Van Hove Singularities 405
12.3 The Thermodynamics of Vibrating Lattices 409
12.3.1 The Ground State of the Lattice and Melting 410
12.3.2 The Specific Heat of the Phonon Gas 413
12.3.3 The Equation of State of the Crystal 418
12.4 Anharmonicity 421
12.4.1 Higher-Order Expansion of the Potential 421
12.4.2 Interaction Among the Phonons 423
12.4.3 Thermal Expansion and Thermal Conductivity in Crystals 425

13 The Experimental Study of Phonons 429
13.1 General Considerations 429
13.2 Optical Methods in the Study of Phonons 431
13.2.1 Infrared Absorption 431
13.2.2 Raman Scattering 433

13.2.3 Brillouin Scattering . . . 436
13.3 Neutron Scattering on a Thermally Vibrating Crystal . . . 438
13.3.1 Coherent Scattering Cross Section . . . 439
13.3.2 Temperature Dependence of the Intensity of Bragg Peaks . . . 443
13.3.3 Inelastic Phonon Peaks . . . 444
13.3.4 The Finite Width of Phonon Peaks . . . 446

14 Magnetically Ordered Systems . . . 449
14.1 Magnetic Materials . . . 450
14.1.1 Ferromagnetic Materials . . . 450
14.1.2 Antiferromagnetic Materials . . . 453
14.1.3 Spiral Magnetic Structures . . . 459
14.1.4 Ferrimagnetic Materials . . . 461
14.2 Exchange Interactions . . . 463
14.2.1 Direct Exchange . . . 463
14.2.2 Indirect Exchange in Metals . . . 464
14.2.3 Superexchange . . . 466
14.2.4 Double Exchange . . . 468
14.3 Simple Models of Magnetism . . . 469
14.3.1 The Isotropic Heisenberg Model . . . 469
14.3.2 Anisotropic Models . . . 471
14.4 The Mean-Field Approximation . . . 473
14.4.1 The Mean-Field Theory of Ferromagnetism . . . 474
14.4.2 The Mean-Field Theory of Antiferromagnetism . . . 478
14.4.3 The General Description of Two-Sublattice Antiferromagnets . . . 485
14.4.4 The Mean-Field Theory of Ferrimagnetism . . . 487
14.5 The General Description of Magnetic Phase Transitions . . . 488
14.5.1 The Landau Theory of Second-Order Phase Transitions . . . 489
14.5.2 Determination of Possible Magnetic Structures . . . 492
14.5.3 Spatial Inhomogeneities and the Correlation Length . . . 494
14.5.4 Scaling Laws . . . 496
14.5.5 Elimination of Fluctuations and the Renormalization Group . . . 500
14.6 High-Temperature Expansion . . . 503
14.7 Magnetic Anisotropy, Domains . . . 504
14.7.1 A Continuum Model of Magnetic Systems . . . 505
14.7.2 Magnetic Domains . . . 508

15 Elementary Excitations in Magnetic Systems . . . 515
15.1 Classical Spin Waves . . . 516
15.1.1 Ferromagnetic Spin Waves . . . 516
15.1.2 Spin Waves in Antiferromagnets . . . 518

15.2 Quantum Mechanical Treatment of Spin Waves 521
15.2.1 The Quantum Mechanics of Ferromagnetic Spin Waves 521
15.2.2 Magnons as Elementary Excitations 524
15.2.3 Thermodynamics of the Gas of Magnons 527
15.2.4 Rigorous Representations of Spin Operators 530
15.2.5 Interactions Between Magnons 533
15.2.6 Two-Magnon Bound States 536
15.3 Antiferromagnetic Magnons 540
15.3.1 Diagonalization of the Hamiltonian 541
15.3.2 The Antiferromagnetic Ground State 543
15.3.3 Antiferromagnetic Magnons at Finite Temperature . . . 544
15.3.4 Excitations in Anisotropic Antiferromagnets 545
15.3.5 Magnons in Ferrimagnets 546
15.4 Experimental Study of Magnetic Excitations 547
15.5 Low-Dimensional Magnetic Systems 548
15.5.1 Destruction of Magnetic Order by Thermal and Quantum Fluctuations 549
15.5.2 Vortices in the Two-Dimensional Planar Model 551
15.5.3 The Spin-1/2 Anisotropic Ferromagnetic Heisenberg Chain 560
15.5.4 The Ground State of the Antiferromagnetic Chain 566
15.5.5 Spinon Excitations in the Antiferromagnetic Chain . . . 569
15.5.6 The One-Dimensional XY Model 572
15.5.7 The Role of Next-Nearest-Neighbor Interactions 575
15.5.8 Excitations in the Spin-One Heisenberg Chain 578
15.5.9 Spin Ladders 581
15.5.10 Physical Realizations of Spin Chains and Spin Ladders 583
15.6 Spin Liquids 584

A Physical Constants and Units 587
A.1 Physical Constants 587
A.2 Relationships Among Units 588

B The Periodic Table of Elements 593
B.1 The Electron and Crystal Structures of Elements 593
B.2 Characteristic Temperatures of the Elements 596

C Mathematical Formulas 601
C.1 Fourier Transforms 601
C.1.1 Fourier Transform of Continuous Functions 601
C.1.2 Fourier Transform of Functions Defined at Lattice Points 606
C.1.3 Fourier Transform of Some Simple Functions 608

C.2 Some Useful Integrals 610
C.2.1 Integrals Containing Exponential Functions 610
C.2.2 Integrals Containing the Bose Function 611
C.2.3 Integrals Containing the Fermi Function 612
C.2.4 Integrals over the Fermi Sphere 613
C.2.5 d-Dimensional Integrals 615
C.3 Special Functions 615
C.3.1 The Dirac Delta Function 615
C.3.2 Zeta and Gamma Functions 617
C.3.3 Bessel Functions 620
C.4 Orthogonal Polynomials 623
C.4.1 Hermite Polynomials 623
C.4.2 Laguerre Polynomials 624
C.4.3 Legendre Polynomials 625
C.4.4 Spherical Harmonics 627
C.4.5 Expansion in Spherical Harmonics 629

D Fundamentals of Group Theory 633
D.1 Basic Notions of Group Theory 633
D.1.1 Definition of Groups 633
D.1.2 Conjugate Elements and Conjugacy Classes 635
D.1.3 Representations and Characters 635
D.1.4 Reducible and Irreducible Representations 637
D.1.5 The Reduction of Reducible Representations 638
D.1.6 Compatibility Condition 639
D.1.7 Basis Functions of the Representations 639
D.1.8 The Double Group 641
D.1.9 Continuous Groups 642
D.2 Applications of Group Theory 646
D.2.1 Irreducible Representations of the Group O_h 646
D.2.2 Group Theory and Quantum Mechanics 648

E Scattering of Particles by Solids 653
E.1 The Scattering Cross Section 653
E.2 The Van Hove Formula for Cross Section 656
E.2.1 Potential Scattering 656
E.2.2 Magnetic Scattering 660

F The Algebra of Angular-Momentum and Spin Operators . . 665
F.1 Angular Momentum 665
F.1.1 Angular Momentum and the Rotation Group 665
F.1.2 The Irreducible Representations of the Rotation Group 667
F.1.3 Orbital Angular Momentum and Spin 669
F.1.4 Addition Theorem for Angular Momenta 670

F.2 Orbital Angular Momentum 672
F.3 Spin Operators 673
F.3.1 Two-Dimensional Representations of the Rotation Group 673
F.3.2 Spin Algebra 674
F.3.3 Projection Operators 676

Figure Credits 677

Name Index 679

Subject Index 683

1

Introduction

It is our everyday experience that materials in the solid, liquid, and gas phases are equally present in the world around us.[1] Nevertheless it is the physics of solids that has acquired a privileged status in the vast field of research into the physical properties of macroscopic bodies. The reason for this is that while investigating liquids and real gases, let alone materials in so-called mesomorphic phases can yield a lot of new and interesting results, it is the processes taking place in solids that have, in the first place, led to a great many fundamentally new concepts. Moreover, solids have a much larger scope of application.

Certain physical properties of solids, most notably the external regularity of crystals have long been known. By the end of the nineteenth century a considerable body of classical knowledge had been amassed about the elastic, thermal, electric, optic, and magnetic properties as well as the symmetries of crystals – without an explanation based on the structure of matter. The birth of solid-state physics can be dated to 1900, when – three years after the discovery of the electron[2] – P. Drude put forward a simple model (based on the late nineteenth century results of statistical mechanics) for the microscopic description of the properties of metals. Using the results of the classical kinetic

[1] This would not be the case in a colder world, as at sufficiently low temperatures and at the same atmospheric pressure all matter would be in the solid phase – with the sole exception of helium. The behavior of the latter is governed by quantum fluctuations, since owing to the small mass of helium atoms these become more important than the weak forces between noble-gas atoms. That is why helium will stay in the liquid phase under atmospheric pressure. To solidify it, the pressure has to reach 25 atm (2.5 MPa) even at low temperatures. The quantum fluid nature of helium manifests itself in yet another way: at very low temperatures liquid ^{3}He and ^{4}He show strikingly different behavior as one is made up of fermionic and the other of bosonic atoms.

[2] The discoverer of the electron, Joseph John Thomson (1856–1940) was awarded the Nobel Prize in 1906 "in recognition of the great merits of his theoretical and experimental investigations on the conduction of electricity by gases."

theory of gases he showed that certain properties of metals can be understood, at least qualitatively, by assuming that electrons move like quasi-free classical particles: billiard balls that collide with obstacles from time to time but otherwise move freely. This picture was developed further by H. A. Lorentz,[3] who gave a somewhat more precise description of the conduction properties of metals in 1905.

The initial successes were soon followed by the first experimental results clearly indicating that certain conduction phenomena could not be properly interpreted within the framework of the Drude–Lorentz model. The most spectacular of these was the discovery of superconductivity by H. Kamerlingh Onnes[4] in 1911.

At the same time the understanding of the structure of solids was advanced by several important discoveries. In 1912 M. von Laue,[5] W. Friedrich, and P. Knipping showed that X-rays (also known as Röntgen rays)[6] diffracted by crystals produce the same interference pattern as light diffracted by an optical grating. The following year W. H. Bragg and W. L. Bragg (father and son)[7] demonstrated that not only the regularity in the atomic positions – i.e., the crystal structure – but also the crystal lattice parameters can be inferred from the interference pattern. They embarked upon the systematic investigation of crystal structures using X-ray diffraction, laying the foundations of the study of crystal structures: crystallography.

Structural studies continue to be an important element of solid-state physical investigations to date; as we shall see later, physical (mechanical, electric, or magnetic) properties are in many respects determined by the structure. However, a large number of phenomena observed in solids – most notably those determined by the behavior of electrons – are less sensitive to structure. The very difference between classical materials science and solid-state physics (in its customary sense) is that the former focuses on applications and therefore does not deal with the properties of materials with an idealized structure; instead it is concerned with the study of how physical properties depend on

[3] Hendrik Antoon Lorentz (1853–1928) and Pieter Zeeman (1865–1943) were awarded the Nobel Prize in 1902, "in recognition of the extraordinary service they rendered by their researches into the influence of magnetism upon radiation phenomena".

[4] Heike Kamerlingh Onnes (1853–1926) was awarded the Nobel Prize in 1913 "for his investigations on the properties of matter at low temperatures which led, inter alia, to the production of liquid helium".

[5] Max von Laue (1879–1960) was awarded the Nobel Prize in 1914 "for his discovery of the diffraction of X-rays by crystals".

[6] Wilhelm Conrad Röntgen (1845–1923) was the first Nobel Prize Winner in physics, in 1901, "in recognition of the extraordinary services he has rendered by the discovery of the remarkable rays subsequently named after him".

[7] Sir William Henry Bragg (1862–1942) and William Lawrence Bragg (1890–1971) were awarded the Nobel Prize in 1915 "for their services in the analysis of crystal structure by means of X-rays".

the real structure. Solid-state physics, on the other hand, is primarily concerned with the interpretation of phenomena, above all those determined by the electrons in solids. The connection between these disciplines is nonetheless strong. One cannot ignore the structures formed by atoms in solid-state physics either since it determines the state of the electrons. We shall see that some of the recently observed most interesting phenomena can manifest themselves only in materials featuring special structures. As it was shown by the example of quasicrystals,[8] the interpretation of the latest discoveries in solid structures present serious challenges for solid-state physicists, too.

The advent of quantum mechanics brought about dramatic changes in the evolution of solid-state physics. It was probably in this field that the new theory had its most spectacular successes, topping the correct qualitative explanations with quantitative ones for a wide range of phenomena. By the second half of the 1920s it had become clear that the state of electrons within solids had to be described using Fermi–Dirac statistics. It then took just about ten years to lay the theoretical foundations that solid-state physics continues to be built upon even today. The most important contributions were due to H. Bethe, F. Bloch, L. Brillouin, W. Heisenberg, L. D. Landau, W. Pauli, J. C. Slater, A. Sommerfeld, A. H. Wilson, and E. P. Wigner.[9]

The forces holding solids together were finally understood in this classical era of solid-state physics through the description provided by quantum mechanics. This allowed a more precise formulation of the vibrations of crystal lattices, and thus the explanation of thermal properties in crystals, the interpretation of conduction and optical properties through the quantum mechanical treatment of electronic states, and, after the identification of the exchange interaction, the elaboration of the theory of magnetic phenomena.

A new generation, including J. Bardeen,[10] R. E. Peierls, and F. Seitz, to name but a few outstanding figures, started to work in the 1930s. During this period, the main lines of research were the experimental and theoretical studies of the properties of metals and insulators. At that time, following the development of quantum mechanics, the theory of metals meant the application of the one-electron approximation – that is, ignoring interactions among electrons or incorporating them into an average potential. While this proved sufficient in many cases, the self-consistent treatment of the average potential necessitated the development of more and more complicated approximation methods, most of which could be treated only numerically.

[8] D. Shechtman, I. Blech, D. Gratias, and J. W. Cahn, 1984.

[9] Among them the following received the Nobel Prize: Hans Albrecht Bethe (1906–2005) in 1967, Felix Bloch (1905–1983) in 1952, Werner Karl Heisenberg (1901–1976) in 1932, Lev Davidovich Landau (1908–1968) in 1962, Wolfgang Pauli (1900–1958) in 1945, and Eugene Paul Wigner (1902–1995) in 1963, although some of the prizes were awarded for achievements in other fields of physics.

[10] John Bardeen was awarded two Nobel Prizes, in 1956 and in 1972, see later.

It is in fact surprising that the one-particle approximation can ever be used, as the system of electrons is a typical many-particle system with a by no means weak interaction, the Coulomb repulsion. That is why field theoretical methods developed in quantum electrodynamics (QED) and quantum field theory (QFT) found new applications in statistical and solid-state physics after the Second World War. This marked the start of the epoch of modern solid-state physics. The application of the methods used in many-body problems to solid-state physics – as pioneered by the *Landau school* and in particular A. A. Abrikosov,[11] L. P. Gorkov, and I. E. Dzyaloshinsky – had a deep impact on the theory of metals, as it provided a consistent approximation scheme for taking the interactions among electrons into consideration.

After the early experimental observation of superconductivity, the exciting problem of working out its microscopic theory remained unsolved for several years. Finally, not only the interaction responsible for superconductivity was identified, but also the theoretical description of the superconducting state was established in 1957 by J. Bardeen, L. N. Cooper, and J. R. Schrieffer.[12] This gave a tremendous boost to both experimental and theoretical investigations into superconductivity. For almost a decade this was the hottest research topic for the solid-state physics community, being the first successfully explained phenomenon for which the usual one-particle approximation failed to provide an adequate interpretation.

Simultaneously, giant steps were made in the understanding of other phenomena of solid-state physics, and devices stemming from these quickly made their way into our everyday life. In 1947 J. Bardeen and W. H. Brattain invented the point-contact transistor, and shortly afterwards W. B. Shockley developed the junction transistor.[13] Yet another branch of solid-state physics burst into blooms: the physics of semiconductors. As a result of its breathtaking development, it has become one of the most important fields of solid-state physics in terms of applications. It is probably in this field that solid-state physics and materials science get closest to each other, since through the step-by-step discovery of new phenomena newer and newer applications may be developed.

In the 1960s research into magnetism gained new momentum as well. Using field theoretical methods, a more precise solution was obtained for the model of magnetism based on localized magnetic moments. At the same time important progress was made in the quantum mechanical treatment of

[11] Alexei Alexeevich Abrikosov (1928–) shared the Nobel Prize with Vitaly Lazarevich Ginzburg (1916–) and Anthony James Leggett (1938–) in 2003 "for pioneering contributions to the theory of superconductors and superfluids".

[12] John Bardeen (1908–1991), Leo Neil Cooper (1930–), and John Robert Schrieffer (1931–) shared the Nobel Prize in 1972 "for their jointly developed theory of superconductivity, usually called the BCS-theory".

[13] William Bradford Shockley (1910–1989), John Bardeen (1908–1991), and Walter Houser Brattain (1902–1987) shared the Nobel Prize in 1956 "for their researches on semiconductors and their discovery of the transistor effect".

magnetism in metals. In 1963, independently of one another, J. Hubbard and M. C. Gutzwiller proposed a seemingly simple model that was expected to give a theoretical description of ferromagnetic behavior caused by non-localized electrons. Although nature is too complex for such an apparently simple theory to yield a valid explanation of magnetism in metals, the Hubbard model and its generalizations nonetheless continue to be the subject of intense research to date.

A breakthrough in the investigations into magnetism came in 1964, when J. Kondo showed in a classic article that the anomalous temperature dependence of resistivity observed in dilute alloys can be explained in terms of the scattering of electrons by magnetic impurities if one goes beyond the customary approximations. This discovery triggered off a veritable avalanche in the experimental and theoretical studies of electronic states around magnetic atoms. It was recognized that some related manifestations of the strong correlations among electrons are difficult to fit into the previous picture of the behavior of fermionic systems, and so new theoretical approaches were called for. The analysis of the problem of magnetic impurities, the so-called Kondo effect then quite naturally led to the correct interpretation of one of the most interesting discoveries of the past decades, the behavior of heavy-fermion systems.

Starting from the 1970s, the experimental methods of solid-state physics have been applied to materials that are not solid in the customary sense of the word, for example polymers and liquid crystals. The discipline encompassing the study of both the customary "hard" materials and such "soft" ones is called *condensed matter physics*. The behavior of crystalline solids on the one hand and polymers and liquid crystals on the other hand share many common points, especially when it comes to phase transitions. To understand critical phenomena the same concepts can be used and the same statistical physical methods may be employed in their quantitative description. As the properties of liquid crystals are not determined by the behavior of electrons but mostly by the geometrical shape of and interactions between large molecules constituting it, they cannot be interpreted along the same lines as those used in the description of the behavior of electrons within solids. Lack of space will prevent us from presenting a discussion of condensed matter physics covering these new aspects as well.

Even in the study of crystalline materials it was a turning point when, during the past decades, the production of newer and newer families of materials, often featuring surprising properties, became possible. A prime example for this was the appearance of organic superconductors in the early 1970s, causing a scientific sensation. In these materials large organic molecules form a highly anisotropic structure in which electrons can propagate more or less freely in one or two directions only. Then in addition to superconductivity, a new type of order, a charge-density-wave state or a spin-density-wave state can also be established. Despite initial hopes, the study of these low-dimensional systems has not provided important new insights into superconductivity, nonetheless

a rich variety of new phenomena has been discovered that cannot be interpreted within the one-particle framework. The reason for this is that correlation among electrons is enhanced because of the spatially restricted character of their motion, so they can give rise to states whose properties are completely different from those observed in ordinary electron systems.

In 1980 K. von Klitzing, G. Dorda, and M. Pepper discovered the quantized Hall effect in a suitably prepared semiconducting structure.[14] Shortly afterwards, the fractional quantum Hall effect was also observed.[15]

The interpretation of these discoveries changed the perspective on the role of impurities and disorder-induced localization, each an interesting field of research in itself. The fractional quantum Hall effect also shed light on a new kind of state in two-dimensional interacting electron gases, enriching the lore of solid-state physics with several new concepts.

In 1986 an experimental observation made by K. A. Müller and J. G. Bednorz[16] sparked off a hitherto unprecedented hunt after superconductors with higher and higher transition temperatures. Although these new superconducting materials have not yet brought a real breakthrough in applications, the fact that their properties are different from those of conventional superconductors has opened new perspectives for considerations about possible novel mechanisms of superconductivity. In this respect, the work of P. W. Anderson[17] is particularly noteworthy. There is still no agreement on what can give rise to superconductivity at such high temperatures but there are more and more signs indicating that the nature of the correlations among electrons is different from that observed in so-called conventional superconductors, and so a theoretical description genuinely different from the Bardeen–Cooper–Schrieffer theory is sought. Research along these lines has taught us a lot about the behavior of strongly correlated electron systems, both in theoretical and experimental aspects.

Solid-state physics experiments are usually conducted on samples that, small as they might be, are macroscopic on atomic scales. Measurements are usually aimed at bulk properties that are independent of the shape and finite

[14] Klaus von Klitzing (1943–) was awarded the Nobel Prize in 1985 "for the discovery of the quantized Hall effect".

[15] D. C. Tsui, H. L. Störmer, and A. C. Gossard, 1982. "For their discovery of a new form of quantum fluid with fractionally charged excitations", Horst Ludwig Störmer (1949–) and Daniel Chee Tsui (1939–) were awarded the Nobel Prize in 1998 together with Robert Betts Laughlin (1950–), who gave the theoretical description of the phenomenon.

[16] Johannes Georg Bednorz (1950–) and Karl Alexander Müller (1927–) shared the Nobel Prize in 1987 "for their important break-through in the discovery of superconductivity in ceramic materials".

[17] Philip Warren Anderson (1923–) shared the Nobel Prize with Sir Nevill Francis Mott (1905–1996) and John Hasbrouck Van Vleck (1899–1980) in 1977 "for their fundamental theoretical investigations of the electronic structure of magnetic and disordered systems".

extent of the sample. The theoretical description is also much simpler when surface phenomena and finite-size effects are ignored. In many cases, however, one's attention turns precisely toward the properties determined by the surface or the finite size of the sample. This is how new fields, such as surface physics, the physics of thin films, and more recently the physics of mesoscopic systems and the physics of nanostructures were born. The extremely fine technologies developed by the semiconducting industry permit the high-precision preparation of samples whose linear dimensions are on the order of 10–100 nm, that is, 100–1000 times larger than atomic distances. In these samples one may study phenomena that occur on scales not much larger than atomic dimensions.

The study of disordered, amorphous systems, glass-like materials, and in particular amorphous semiconductors and metallic glasses has recently received an increasing attention in solid-state physics, too. Materials science had been interested in such materials for some time because of their practical applicability, however, understanding their behavior is equally important in solid-state physics and in statistical physics. This is especially true for spin glasses, whose theoretical description has required the introduction of new concepts and the development of novel theoretical methods. These concepts and method have been found to be applicable to phenomena and systems beyond the traditional scope of physics, such as stock markets or behavior research. The term *physics of complex systems* is used for the discipline where the methods of statistical and solid-state physics are applied to such new fields.

In this three-volume treatise the presentation of solid-state physics follows the historical development outlined above only in the sense that we shall encounter newer and newer phenomena and will be led to more and more complex considerations. Our primary aim is to show how one can determine the properties of solids using the methods of quantum mechanics – basing the discussion, as much as possible, on first principles –, and how one can interpret the observed behavior of solids. However, solid-state physics is a science that is both experimental and theoretical, with the characteristic features of both approaches. Therefore besides theoretical explanations, one should always be aware of the experimental methods for investigating the discussed phenomena. Some of these techniques are extremely simple, accessible in practically all solid-state physics laboratories, while others require state-of-the-art technology or large-scale equipment. In this book we shall indicate at the appropriate places how one may study certain phenomena, and on some occasions we shall present experimental methods in some detail.

The first volume begins with a brief introduction into the structure of condensed matter; then some simple properties – known from classical physics, atomic physics, or statistical physics – of the building blocks of solids (ion cores) are recalled. The discussion of forces that hold solids together in the condensed phase is followed by the presentation of the structure determined by atomic positions within the solid, its defects, and the dynamical properties

of the structure. Formal analogy justifies treating magnetically ordered structures and some simple questions of their dynamics in the same volume.

In the second volume we turn to our main task, the study of the states and behavior of the system of electrons. Its importance is clearly shown by the fact that among the mechanical, thermal, electric, optical, and magnetic properties of solids all except for the first and to some extent the second group are primarily determined by the behavior of electrons. We shall start out with the discussion of a gas of free electrons and then gradually take into account their interactions with atoms. First we shall be concerned only with the effects of the periodic potential of atoms sitting statically in a crystal lattice, and only later shall we examine how the behavior of the electrons is affected by the oscillatory motion of the atoms. After the presentation of both theoretical and experimental methods for determining the electronic state and the semiclassical treatment of the dynamics of electrons we shall devote separate chapters to the properties of metals, semiconductors, insulators, and superconductors. Throughout this volume, we shall use the one-particle approximation.

A more profound study of the interaction among electrons is presented in the third volume. The analysis of the correlations among electrons will lead us to the instabilities occurring in the electron gas that are responsible for the appearance of magnetic, superconducting, charge-density-wave and spin-density-wave states. The microscopic theory of superconductivity is also discussed there. Finally, the effects of strong electron correlations are explored, and some questions of the physics of disordered systems are addressed.

A series of appendices conclude each volume. The first appendix of the present volume contains the numerical values of fundamental physical constants. The next covers some properties of the elements in the periodic table that play important roles in solid-state physics. The one on mathematical relations provides a summary of the conventions used in Fourier transformations, some useful integrals, as well as the essentials about the special functions used in the text. This is followed by a summary of group theory, the scattering of particles by solids, and the quantum theory of spin and orbital angular momentum. The fundamentals of many-body problems are presented in the appendices of the second and third volumes. None of the appendices purport to be complete, they just evoke the basic concepts that the reader should be familiar with to be able to follow the arguments of the text.

Further Reading

To complement the material in the present book, the interested reader can consult a wide range of solid-state physics textbooks. An exhaustive listing would be impossible, therefore a rather subjective selection is given below that contains a few classics and some newer texts.[18] Some treat the subject

[18] For some classic texts that have been republished in unaltered form the original year of publishing is given. Otherwise the year of the last edition is usually given.

at an introductory level, while others are more advanced and can serve as useful references for undergraduate or graduate students preparing for their final exams.

Introductory textbooks

1. N. W. Ashcroft and N. D. Mermin, *Solid State Physics*, Holt, Rinehart and Winston, New York (1976).
2. G. Burns, *Solid State Physics*, Academic Press, Inc., Orlando, Florida (1990).
3. P. M. Chaikin and T. C. Lubensky, *Principles of Condensed Matter Physics*, Cambridge University Press, Cambridge (1995).
4. J. R. Christman, *Fundamentals of Solid State Physics*, John Wiley & Sons, New York (1988).
5. J. R. Hook and H. E. Hall, *Solid State Physics*, Second Edition, John Wiley & Sons, Chichester (1991).
6. H. Ibach and H. Lüth, *Solid-State Physics, An Introduction to Principles of Materials Science*, Third edition, Springer-Verlag, Berlin (2003).
7. C. Kittel, *Introduction to Solid State Physics*, Eighth edition, John Wiley & Sons, New York (2004).
8. R. Kubo and T. Nagamiya, *Solid State Physics*, McGraw-Hill Book Co., Inc., New York (1969).
9. M. P. Marder, *Condensed Matter Physics*, 5th corrected printing, John Wiley & Sons, Inc., New York (2004).
10. U. Mizutani, *Introduction to the Electron Theory of Metals*, Cambridge University Press, Cambridge (2001).
11. E. Mooser, *Introduction à la physique des solides*, Presses polytechniques et universitaires romandes, Lausanne (1993).
12. H. P. Myers, *Introductory Solid State Physics*, Second Edition, Taylor & Francis, London (1997).
13. Oxford Master Series in Condensed Matter Physics, Oxford University Press, Oxford.
    – J. Singleton, *Band Theory and Electronic Properties of Solids* (2001).
    – A. M. Fox, *Optical Properties of Solids* (2001).
    – S. Blundell, *Magnetism in Condensed Matter* (2001).
    – R. A. L. Jones, *Soft Condensed Matter* (2002).
    – M. T. Dove, *Structure and Dynamics, An Atomic View of Materials* (2003).
    – J. F. Annett, *Superconductivity, Superfluidity and Condensates* (2004).
14. H. M. Rosenberg, *The Solid State*, Third Edition, Oxford Physics Series, Oxford University Press, Oxford (1988).

15. U. Rössler, *Solid State Theory, An Introduction*, Advanced Texts in Physics, Springer-Verlag, Berlin (2004).
16. R. J. Turton, *The Physics of Solids*, Oxford University Press, Oxford (2000).

Textbooks with a more theoretical approach

1. P. W. Anderson, *Concepts in Solids*, The Benjamin-Cummings Publishing Co., Inc., Reading, Massachusetts (1963).
2. P. W. Anderson, *Basic Notions of Condensed Matter Physics*, The Benjamin-Cummings Publishing Co., Inc., Menlo Park, California (1984).
3. A. O. E. Animalu, *Intermediate Quantum Theory of Crystalline Solids*, Prentice Hall, Inc., Englewood Cliffs, New Jersey (1977).
4. J. Callaway, *Quantum Theory of the Solid State*, Second Edition, Academic Press, Inc., Boston (1991).
5. W. A. Harrison, *Solid State Theory*, McGraw-Hill Book Co., Inc. New York (1970).
6. W. Jones and N. H. March, *Theoretical Solid State Physics*, Wiley-Interscience, New York (1973).
7. C. Kittel, *Quantum Theory of Solids*, 2nd revised edition, John Wiley & Sons, New York (1987).
8. O. Madelung, *Introduction to Solid-State Theory*, Springer Series in Solid-State Sciences, 3rd printing, Springer-Verlag, Berlin (1996).
9. P. Phillips, *Advanced Solid State Physics*, Westview Press, Boulder (2002).
10. P. L. Taylor and O. Heinonen, *A Quantum Approach to Condensed Matter Physics*, Cambridge University Press, Cambridge (2002).
11. *The Structure and Properties of Matter*, Editor: T. Matsubara, Springer-Verlag, Berlin (1982).
12. S. V. Vonsovsky and M. I. Katsnelson, *Quantum Solid-State Physics*, Springer-Verlag, Berlin (1989).
13. J. M. Ziman, *Principles of the Theory of Solids*, Cambridge University Press, Cambridge (1979).

Review articles and monographs

The interested reader is referred to the volumes of the following series which are not textbooks but review articles and monographs.

1. *Solid State Physics, Advances in Research and Applications*, Founding Editors: F. Seitz and D. Turnbull, Editors: H. Ehrenreich and F. Spaepen, Academic Press, San Diego.

2. *Springer Series in Solid-State Sciences*, Edited by P. Fulde, Springer-Verlag, Berlin.

Collections of problems in solid-state physics

Instructors may use certain sections of the book at tutorials but cannot assign the discussed problems to students as homework since their solutions are also given. Two good collections of further solid-state physics problems are:

1. L. Mihály and M. C. Martin, *Solid State Physics; Problems and Solutions*, John Wiley & Sons, Inc., New York (1996).
2. *Problems in Solid State Physics*, Editor: H. J. Goldsmid, Academic Press, New York (1968).

Regularly updated handbooks presenting physical constants and experimental data

1. *Landolt–Börnstein, Numerical Data and Functional Relationships in Science and Technology*, New Series, Editor in Chief: O. Madelung, Springer-Verlag, Berlin.

   Some of the data and the diagrams are available on the internet at the Springer website http://www.springer.com.
2. *CRC Handbook of Chemistry and Physics*, Editor-in-Chief: D. R. Lide, CRC Press, Boca Raton.
3. *Springer Handbook of Condensed Matter and Materials Data*, Editors: W. Martienssen and H. Warlimont, Springer-Verlag, Berlin (2005).

2

The Structure of Condensed Matter

One of the most characteristic features of solids is their relatively high mechanical rigidity, that is, resistance to external forces that would force them to change their shape. At first sight this is what distinguishes them from matter in the liquid and gas phases. The word "solid" refers to this very property. However, this rigidity is not perfect. Weaker forces will produce elastic strain, while stronger ones may cause plastic deformation or rupture.

Mechanical rigidity is due to stronger or weaker bonds that hold atoms or molecules together in a solid. A solid can be classically pictured as a collection of atoms held together by springs. These springs tend to hinder the free displacement of atoms relative to each other. Because of thermal motion, however, the atoms will not be strictly at rest but will oscillate about their equilibrium position within the cavity among its neighbors. At sufficiently high temperatures the oscillation amplitude – and thus the mean square displacement from equilibrium – can become so large that the atoms are no longer localized. This corresponds to the melting of the solid.[1] When temperature is increased even further, thermal motion completely overcomes binding forces, and the liquid vaporizes.

In substances built up of large nonspherical molecules transition from solid to liquid phase may not occur in a single step but through intermediate, so-called *mesomorphic* phases. In such phases substances are less rigid than in their solid phase; in many respects they are closer to liquids. A clear distinction from gases is offered by the collective term *condensed phases* for the solid, mesomorphic, and liquid phases.

In this chapter we shall present the general characteristics of condensed matter, paying special attention to those nonsolid phases that we shall not discuss in detail elsewhere.

[1] The melting point of the elements is listed in Appendix B.

2.1 Characterization of the Structure

The arrangement of atoms is usually not completely random in condensed phases. Interatomic forces create some order in the atomic structure. This order may apply to the entire sample or it may manifest itself only locally; it may comprise the internal atomic degrees of freedom but it may just as well be restricted to atomic positions.

2.1.1 Short- and Long-Range Order

The largest and by far most thoroughly analyzed class of solids is that of *crystalline solids*. In *ideal crystals* atoms are arranged in strictly periodic arrays, i.e., spaced at regular distances in each direction. This perfect order of the crystalline structure is assumed to be maintained indefinitely. The system is homogeneous in the sense that the neighborhoods of any two equivalent atoms are identical no matter how far they are separated. The crystal is then said to possess *long-range order*. In such cases atomic densities show strong correlations even for large spatial and temporal separations. At finite (nonzero) temperatures positions are always slightly smeared out by thermal motion, thus correlations are reduced but they are not destroyed. Long-range order is preserved in the crystal up to the melting point.

A large proportion of the substances do not exhibit long-range order in the condensed phase. Atoms separated by macroscopic distances are uncorrelated in such samples. In other cases long-range order is present in certain spatial directions but absent in others. Nonetheless atomic positions may still be related (correlated) over microscopic distances comparable to atomic dimensions. The appearance of such correlations is most readily understood in substances where covalent bonds play an essential role. Because of the directionality of these bonds, in all parts of the sample the relative orientation of the first few neighbors, located at more or less regular distances from each other, is fairly definite. Locally, over distances among the first few neighbors (in other words, on scales comparable to atomic dimensions) some kind of order is observed that is similar to the one in crystalline solids. Over larger distances, deviations from the regular bond directions and distances may become so important that correlations among atomic positions are lost. In such cases one speaks of *short-range order*.

Short- and long-range order can be observed not only in the spatial arrangement of atoms but also in their internal degrees of freedom, for example the orientation of their magnetic moments. In Chapter 14 we shall study magnetic systems in which atomic magnetic moments exhibit long-range order. Below we shall consider systems built up of identical atoms and we shall be concerned with their spatial arrangement only. These considerations will be generalized in Chapter 10 to multicomponent substances and in Chapter 28 to systems of electrons, where will shall also analyze the short-range order in the spatial arrangement of spins.

Owing to the long-range order in the atomic arrangement, structure can be completely characterized by a handful of parameters in crystalline materials. In other cases even if the position and the expectation values of internal degrees of freedom were known for the very large number of atoms, specifying them would give an inextricable and unmanageable set of data. Instead of this the structure can be fully characterized in terms of atomic distribution functions.[2]

The n-particle probability density function $\rho_n(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_n)$ is defined by stipulating that the probability that the volumes $\mathrm{d}\boldsymbol{r}_1$ around $\boldsymbol{r}_1$, $\mathrm{d}\boldsymbol{r}_2$ around $\boldsymbol{r}_2$, etc. contain precisely one atom each be given by

$$\mathrm{d}P_n(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_n) = \rho_n(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_n)\, \mathrm{d}\boldsymbol{r}_1\, \mathrm{d}\boldsymbol{r}_2 \ldots \mathrm{d}\boldsymbol{r}_n\,. \tag{2.1.1}$$

In disordered systems where a large number of random atomic configurations are possible, the n-particle probability density function is determined by averaging over all possible configurations. At finite (nonzero) temperatures, where particles are in thermal motion, a so-called thermal average is taken in which each possible state is multiplied by an energy-dependent statistical weight, the Boltzmann factor. The n-particle probability distribution function is obtained by dividing the n-particle density function by the one-particle densities:

$$g_n(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_n) = \frac{\rho_n(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_n)}{\rho_1(\boldsymbol{r}_1)\rho_1(\boldsymbol{r}_2) \ldots \rho_1(\boldsymbol{r}_n)}\,. \tag{2.1.2}$$

For a complete description of a disordered physical system, the infinite hierarchy of such expressions would be required in principle. However, we shall see later that it is sufficient to know the one- and two-particle distribution functions – which can be determined from experiments. In what follows we shall study only these.

Consider a system with N atoms in a volume V, and denote the position vector of the ith atom by $\boldsymbol{R}_i$. The one-particle probability density function, ρ_1, which we shall call ρ, is

$$\rho(\boldsymbol{r}) = \left\langle \sum_{i=1}^{N} \delta(\boldsymbol{r} - \boldsymbol{R}_i) \right\rangle, \tag{2.1.3}$$

where $\langle \ldots \rangle$ denotes the configurational or thermal average. The above expression is the actual density. The two-particle probability density function ρ_2, which we shall henceforth call P, is given by

$$P(\boldsymbol{r}_1, \boldsymbol{r}_2) = \left\langle \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} \delta(\boldsymbol{r}_1 - \boldsymbol{R}_i)\delta(\boldsymbol{r}_2 - \boldsymbol{R}_j) \right\rangle. \tag{2.1.4}$$

[2] Later we shall introduce further functions that indicate whether the positions of atoms at different times are correlated. Here we shall consider the static case only.

The definition of probability densities implies

$$\int_V \rho(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r} = N, \tag{2.1.5}$$

$$\int_V P(\boldsymbol{r}_1,\boldsymbol{r}_2)\,\mathrm{d}\boldsymbol{r}_2 = (N-1)\rho(\boldsymbol{r}_1)\,, \tag{2.1.6}$$

and, of course,

$$\iint_V P(\boldsymbol{r}_1,\boldsymbol{r}_2)\,\mathrm{d}\boldsymbol{r}_1\,\mathrm{d}\boldsymbol{r}_2 = N(N-1)\,. \tag{2.1.7}$$

If long-range order is present, the positions of the two atoms are correlated even if the separation between $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ is very large. If, however, there is no long-range order then the correlation between atomic positions is washed out at large separations, and so

$$P(\boldsymbol{r}_1,\boldsymbol{r}_2) \to \rho(\boldsymbol{r}_1)\rho(\boldsymbol{r}_2)\,, \qquad \text{if} \quad |\boldsymbol{r}_1-\boldsymbol{r}_2| \to \infty\,. \tag{2.1.8}$$

One separates out this expression to define the *correlation function* $C(\boldsymbol{r}_1,\boldsymbol{r}_2)$:

$$C(\boldsymbol{r}_1,\boldsymbol{r}_2) = P(\boldsymbol{r}_1,\boldsymbol{r}_2) - \rho(\boldsymbol{r}_1)\rho(\boldsymbol{r}_2)\,. \tag{2.1.9}$$

This correlation function indeed indicates whether the presence of an atom at $\boldsymbol{r}_1$ affects the probability of finding another atom at $\boldsymbol{r}_2$. For perfectly random atomic arrangements the correlation function is identically zero. For amorphous systems with short-range order the function takes finite values at small separations and drops off exponentially at large distances. On the other hand, for crystalline samples the function shows the same periodicity as the underlying structure even at large separations.

We shall focus on the distribution function rather than the probability density function and suppress the index in the highly important two-particle expression:

$$g(\boldsymbol{r}_1,\boldsymbol{r}_2) = \frac{P(\boldsymbol{r}_1,\boldsymbol{r}_2)}{\rho(\boldsymbol{r}_1)\rho(\boldsymbol{r}_2)}\,. \tag{2.1.10}$$

In what follows, we shall almost exclusively study homogeneous systems, in which the one-particle density is uniform, and we shall denote the ratio N/V by n. The two-particle probability density will depend only on the difference of $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$; this is clearly seen when P is written as

$$\begin{aligned} P(\boldsymbol{r}_1,\boldsymbol{r}_2) &= \frac{1}{V}\left\langle \int_V \sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N} \delta(\boldsymbol{r}_1-\boldsymbol{r}'-\boldsymbol{R}_i)\delta(\boldsymbol{r}_2-\boldsymbol{r}'-\boldsymbol{R}_j)\,\mathrm{d}\boldsymbol{r}' \right\rangle \\ &= \frac{1}{V}\left\langle \sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N} \delta(\boldsymbol{r}_1-\boldsymbol{r}_2-\boldsymbol{R}_i+\boldsymbol{R}_j) \right\rangle. \end{aligned} \tag{2.1.11}$$

In terms of the new variable $\boldsymbol{r} = \boldsymbol{r}_1 - \boldsymbol{r}_2$ we have

$$P(\boldsymbol{r}) = \frac{1}{V} \left\langle \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} \delta(\boldsymbol{r} - \boldsymbol{R}_i + \boldsymbol{R}_j) \right\rangle , \tag{2.1.12}$$

and the two-particle distribution function becomes

$$g(\boldsymbol{r}) = \frac{V}{N^2} \left\langle \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} \delta(\boldsymbol{r} - \boldsymbol{R}_i + \boldsymbol{R}_j) \right\rangle . \tag{2.1.13}$$

In isotropic systems the distribution function $g(\boldsymbol{r})$ depends only on $r = |\boldsymbol{r}|$. The quantity $g(r)$ is then called the *radial distribution function*. If an atom is now selected, the average number of particles within a spherical shell of radius r and thickness $\mathrm{d}r$ around it is

$$n(r)\,\mathrm{d}r = 4\pi n\, g(r)\, r^2\, \mathrm{d}r\,, \tag{2.1.14}$$

which explains why sometimes $4\pi r^2 g(r)$ rather than $g(r)$ is called the radial distribution function.

If there is only short-range order then $g(\boldsymbol{r}) \to 1$ for $|\boldsymbol{r}| \to \infty$. When this term is separated out, one is left with the *pair-correlation function*

$$c(\boldsymbol{r}) = g(\boldsymbol{r}) - 1\,. \tag{2.1.15}$$

We shall often encounter this correlation function, sometimes in another form that contains $i = j$ terms as well. We therefore introduce the expression

$$\Gamma(\boldsymbol{r}) = \frac{1}{N} \left\langle \sum_{i=1}^{N} \sum_{j=1}^{N} \delta(\boldsymbol{r} - \boldsymbol{R}_i + \boldsymbol{R}_j) \right\rangle . \tag{2.1.16}$$

When the term $i = j$ is separated, it is readily seen that

$$\Gamma(\boldsymbol{r}) = \delta(\boldsymbol{r}) + n\, g(\boldsymbol{r})\,. \tag{2.1.17}$$

By taking the Fourier transform of both sides,

$$\Gamma(\boldsymbol{K}) = 1 + n \int_V g(\boldsymbol{r}) \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}}\, \mathrm{d}\boldsymbol{r}\,. \tag{2.1.18}$$

Separating from $\Gamma(\boldsymbol{K})$ the Fourier component $\boldsymbol{K} = 0$, which is equal to the number of particles regardless of the structure, we have

$$\Gamma(\boldsymbol{K}) = N\delta_{\boldsymbol{K},0} + S(\boldsymbol{K}). \tag{2.1.19}$$

This expression defines the *structure factor* $S(\boldsymbol{K})$.[3] Using the relation

$$n \int_V \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}}\,\mathrm{d}\boldsymbol{r} = N\delta_{\boldsymbol{K},0}\,, \tag{2.1.20}$$

and comparing (2.1.18) and (2.1.19),

$$S(\boldsymbol{K}) = 1 + n\int_V [g(\boldsymbol{r}) - 1]\,\mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}}\,\mathrm{d}\boldsymbol{r}\,. \tag{2.1.21}$$

An inverse Fourier transform now gives

$$c(\boldsymbol{r}) = g(\boldsymbol{r}) - 1 = \frac{1}{(2\pi)^3 n}\int [S(\boldsymbol{K}) - 1]\,\mathrm{e}^{\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}}\,\mathrm{d}\boldsymbol{K}\,. \tag{2.1.22}$$

In the isotropic case angular integration in (2.1.21) leads to

$$S(K) = 1 + n\int_0^\infty 4\pi r^2\,[g(r) - 1]\,\frac{\sin Kr}{Kr}\,\mathrm{d}r\,, \tag{2.1.23}$$

that is, $S(\boldsymbol{K})$ depends only on $K = |\boldsymbol{K}|$. Now the inverse transformation gives

$$c(r) = g(r) - 1 = \frac{1}{8\pi^3 n}\int_0^\infty 4\pi K^2\,[S(K) - 1]\,\frac{\sin Kr}{Kr}\,\mathrm{d}K\,. \tag{2.1.24}$$

The definition (2.1.19) of $S(\boldsymbol{K})$ implies that $S(\boldsymbol{K})$ vanishes for $\boldsymbol{K} = 0$. It can also be seen from the equations that the normalization conditions on the pair-correlation function implies that $S(K)$ vanishes in the $K \to 0$ limit as well. The limiting process must be treated with care when the spatial distribution of the atoms is totally uncorrelated, and thus $g(r) = 1$ follows from $c(r) = 0$. In this case everywhere except for the point $K = 0$ the structure factor is constant, $S(K) = 1$.

As we shall see later, not only does the pair-correlation function lend itself to simple theoretical interpretation, but – with certain restrictions – it can also be determined from measurements. The cross section in elastic scattering experiments, e.g. in X-ray diffraction is proportional to the structure factor $S(\boldsymbol{K})$. From the measured $\boldsymbol{K}$-dependence of the structure factor the spatial correlations among atoms can be inferred. Some examples of the radial distribution function will be presented later.

[3] In the literature the term *structure factor* is sometimes used for related but not identical expressions as well. See the footnote on page 248.

2.1.2 Order in the Center-of-Mass Positions, Orientation, and Chemical Composition

Up to now atoms have been assumed to be point-like, but our considerations are equally valid for condensed matter built up of spherical atoms or molecules. The centers of mass of the building blocks (atoms or molecules) may exhibit short- or long-range order. If a long-range order is present then – as we shall examine it in detail in Chapter 5 – the full translational symmetry of the sample observed macroscopically is broken on atomic scales. On the other hand, if a snapshot is taken of the liquid, the centers of mass are found to be distributed randomly, and so invariance under arbitrary translations is preserved. The generalization of this conclusion is also justified: the appearance of order is always accompanied by the breaking of some symmetry.

Besides order in the positions of the centers of mass, other types of order may also appear in the condensed phase if the building blocks of the sample are not spherical. In substances made up of large molecules, building blocks are often long, rod-shaped, or flat, disk-like molecules. In the condensed phase – almost independently of the ordered or disordered arrangement of the centers of mass – the rod axes or the disk planes may also be ordered. Denoting the angle between a certain reference direction and the rod axis or the normal to the disk plane by θ, one can determine the quantity $\left\langle \cos^2\theta - \frac{1}{3} \right\rangle$, where $\langle \dots \rangle$ denotes the average over all molecules. If this quantity is different from zero then the sample is said to exhibit orientational order.

Orientational order may also be present when the direction is not determined by the form of the molecules but by an internal degree of freedom, e.g. atomic magnetic moment. To characterize magnetic materials the specification of the orientation of magnetic moments is just as necessary as the specification of atomic positions. In ordered magnetic structures where the directions of magnetic moments are also ordered, this orientational order may show an even greater diversity than the crystalline arrangements of atomic positions. We shall return to this problem in Chapter 14 on magnetic materials.

Disorder may also be rooted in chemical composition. In nonstoichiometric alloys even when atoms occupy the sites of a regular lattice, the order is imperfect, as the distribution of the components over the lattice sites is identical only in an average sense in different parts of the sample. The same chemical disorder may also appear in stoichiometric alloys at sufficiently high temperatures, since due to its entropy, a disordered state has lower free energy than an ordered one. Chemical ordering occurs through a disorder–order phase transition, and long-range order appears only below the critical point. Short-range order may nonetheless exist in the high-temperature phase. As an example, consider a material composed of two types of atoms, A and B, and assume that the configuration in which atoms of either type are surrounded by atoms of the opposite type is energetically more favorable than the configuration in which atoms of the same kind are next to each other. Then the majority of the nearest neighbors will be atoms of the opposite type. At small scales the

material seems to be chemically ordered. When temperature is increased, the free energy is more and more dominated by the term $-TS$, and short-range order is gradually destroyed.

2.2 Classification of Condensed Matter According to Structure

One possible classification of condensed matter is according to the degree center-of-mass or orientational order is present. Before turning to the vast subject of the study of crystalline materials, we shall give a highly simplified and very concise overview of the characteristics of the structure of condensed matter.

2.2.1 Solid Phase

When classifying materials that are, from a mechanical point of view, in the solid phase, if small oscillations of the atoms are neglected, the atomic arrangement is found to be regular in some substances and irregular in others. Thus solids are divided into two broad categories according to their structure: *crystalline* and *noncrystalline*. It should be noted that certain mesomorphic phases can also be considered solid from a mechanical viewpoint. However, due to their special structure, they will be considered separately.

Crystalline Solids

Although it is impossible to prove it with mathematical rigor, it is a generally accepted assumption that at low temperatures crystalline structure is the energetically most stable. This statement is based on the experience that it is indeed possible to produce genuinely regular structures using crystal-growth processes – provided they are slow enough and so the system has sufficient time to find the most stable, lowest-energy state among local minima.

In a crystal each atom sits at a well-defined site that is easily determined. This also means that in crystals built up of molecules both the molecular centers of mass and the orientation of the molecular axes with respect to the crystallographic axes show long-range order. In such cases the pair-correlation function is the sum of a periodic sequence of Dirac deltas, and thus in $\boldsymbol{K}$-space $S(\boldsymbol{K})$ will also be the sum of a periodic sequence of Dirac deltas. We shall see this in more detail in Chapter 8.

If the orientation of the crystallographic axes is the same throughout the sample then the sample is called a *single crystal*. In *real crystals*, however, the order is never perfect. First, there may be defects in the atomic arrangement due to imperfect crystal growth, and these defects can destroy the correlation between the positions of distant atoms. For example, if the crystal starts to

grow at several points then the sample will consist of crystal grains of different (usually macroscopic) sizes and irregular shapes, so-called *crystallites*. The crystallographic axes of individual grains are oriented independently of each other. Such samples are said to be *polycrystalline*.

Second, at finite temperatures the ubiquitous thermal fluctuations may also create defects, which will disrupt strict periodicity over large distances. Nonetheless, the physical properties of such materials – and also of polycrystalline samples, provided the crystallites are not too small – are in many respects similar to those of ideal crystals, so our results are usually valid for them as well.

Noncrystalline Solids

With the sole exception of helium, all samples in thermodynamic equilibrium are expected to be in an ordered crystalline phase at low temperatures. However, using quenching (rapid cooling) or other methods, it is equally possible to produce samples that are solid from a mechanical point of view but in which atoms are frozen in random (disordered) positions – unlike in crystals, where they arrange themselves into a regular periodic array. The prime example for this is glass. Therefore solid materials with a perfectly disordered structure are often called *glasses*. For example, in *metallic glasses* two or more metallic components are arranged randomly – but in such a way that atoms are packed closely to fill space as tightly as possible. The structure of such *amorphous materials* is disordered, and so – in contrast to crystals – long-range order is absent. No correlation is found between the positions of two distant atoms; the correlation function defined in (2.1.9) vanishes at large distances. Short-range order may, however, exist, as we shall see it in Chapter 10. An important difference with liquids is that in amorphous solids the disordered atomic positions do not vary with time, that is why these materials are sometimes called *solid solutions* as well.

As it was mentioned in the introductory chapter, 1984 brought the discovery of a new type of solid material, which, in a sense, is halfway between crystalline and disordered systems. There is some sort of long-range order in the spatial orientation of relative atomic positions, however no regular periodic structure is formed. We shall see in Chapter 10 that the pair-correlation function is then similar to that of amorphous materials, while the diffraction pattern and the structure factor $S(K)$ – determined using methods of structural analysis – are similar to those of crystalline materials. For this reason, they are called *quasicrystals*.

In the largest part of our solid-state physics studies we shall be concerned with the properties of crystalline solids, since classical solid-state physics is the physics of crystalline materials. The behavior of noncrystalline materials has nevertheless attracted more and more attention recently. Therefore our investigation into crystalline structures will be followed by the presentation of the most important structural characteristics of such systems (Chapter 10)

and some of their physical properties (Chapter 36). As we shall not discuss the properties of liquid crystals and plastic crystals elsewhere, below we shall give a brief overview of some of the characteristic features of their structure.

2.2.2 Liquid Phase

In liquids no long-range order is present either in atomic (molecular) positions or the relative orientation of atomic (molecular) axes. If a snapshot is taken of the atoms of the liquid at any instant of time, the static two-particle distribution function $g(\boldsymbol{r}_1, \boldsymbol{r}_2)$ is found to be homogeneous and isotropic – in other words, $g(\boldsymbol{r}_1, \boldsymbol{r}_2)$ depends only on $r = |\boldsymbol{r}| = |\boldsymbol{r}_1 - \boldsymbol{r}_2|$ and for large separations it tends to a constant value that is independent of the direction and the separation. This is readily seen in Fig. 2.1, which shows the experimentally determined structure factor $S(K)$ for liquid argon and the two-particle radial distribution function obtained from its Fourier transform.

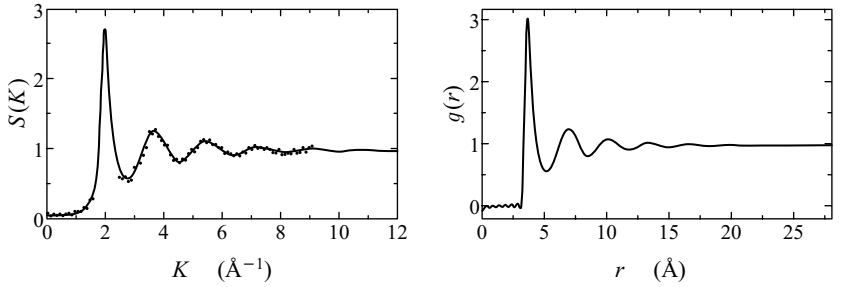


Fig. 2.1. The structure factor $S(K)$ of liquid argon, measured at 85 kelvins in neutron scattering experiments, and the radial distribution function obtained from its Fourier transform [J. L. Yarnell et al., *Phys. Rev. A* **7**, 2130 (1973)]

Similar things would be observed in other liquids. The radial distribution function approaches unity for large distances showing that in liquids there is no long-range correlation among the atoms. On the other hand, the sharp peak in the structure factor $S(K)$ and the oscillation following it – or, in terms of the derived radial distribution function, the small number of relatively sharp peaks at short distances – indicate short-range order among atoms. An explanation for this short-range order is provided by the simplest model for the structure of liquids, the *Bernal model*.[4]

This model, shown in Fig. 2.2, is obtained by arranging the atoms – considered to be rigid spheres – randomly next to each other so that they should be quite closely packed. As it is seen in the magnified part, whichever atomic sphere is selected, due to close packing, the distance to the centers of its nearest neighbors is equal to or just slightly larger than twice the atomic radius.

[4] J. D. Bernal, 1959.

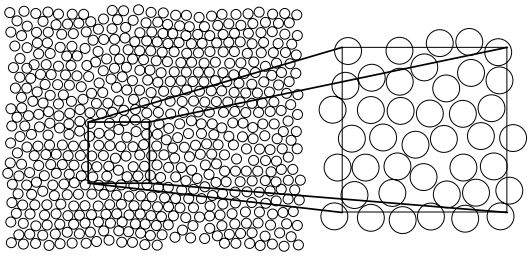


Fig. 2.2. The Bernal model of liquids, with a relatively close-packed random arrangement of atomic spheres on two different length scales

Therefore $g(r)$ is sharply peaked at twice the atomic radius. The system is thus not homogeneous on such small scales. In subsequent layers the distribution of the centers is smoother, and correlation between the positions of atomic centers gets weaker with increasing separation.

Although space is not so tightly filled in real liquids, the radial distribution function indicates strong correlations. From the position of the first maximum $K_{\max}$ in $S(K)$, the distance r_0 to the nearest neighbors is directly calculated from $K_{\max} \approx 2\pi/r_0$, in agreement with the nearest-neighbor distance read off from the graph of $g(r)$. The distribution is less sharply peaked in the second and third layers but the peak is nonetheless clearly visible. For atoms separated by even larger distances there is practically no correlation.

Qualitatively one may say that the sharper and more numerous the peaks, the longer the range of the correlation among atomic positions. The limiting case is that of an ideal crystal where – as mentioned above – the structure factor $S(\boldsymbol{K})$ is a sequence of infinitely sharp discrete peaks (Dirac deltas).

2.2.3 Mesomorphic Phases

It is always possible to draw a sharp line between the order in the centers of mass and in the orientation of molecules/atoms observed in crystals on the one hand, and the disordered atomic arrangement in the molten phase at high temperatures or in the glass-like phase obtained from the melt by quenching on the other hand, in the sense that one can always determine whether or not long-range order is present. In other words, the state with broken translational and rotational symmetry can be clearly distinguished from the state with full translational and rotational symmetry. However, transition between ordered and disordered states does not always happen in a single step. Order in the spatial arrangement of molecular centers of mass and in molecular orientations do not necessarily appear or disappear at the same time, i.e., the breaking of translational and rotational symmetry may occur at different temperatures. This is how mesomorphic phases are formed. When starting from the low-temperature solid phase, if center-of-mass order is disrupted but orientational order is maintained in the phase transition, the new phase is called the *liquid*

crystal phase. In some cases the destruction of the center-of-mass order does not occur simultaneously along the three directions. It might happen along one or two directions only, leading to different types of liquid crystals. On the other hand, if center-of-mass order is maintained in the phase transition and only orientational order is lost in the new phase then one speaks of *plastic crystals*.

Liquid Crystals

The liquid-crystalline phase usually manifests itself as one of the condensed phases of mesogenic compounds (mesogens) made up of long (rod- or lath-like), or flat (disk- or sheet-shaped) molecules – or molecules containing a rigid molecular moiety (mesogenic unit) of one of the above shapes that plays an important role in the formation of the mesomorphic phase.[5] Figure 2.3 shows some classic examples of long organic molecules that exhibit a liquid-crystalline phase. PAA stands for *para*-azoxy anisole, MBBA for 4-methoxy-benzilidene-4-butyl-aniline, and TBBA for terephtal-bis-butyl-aniline. Figure 2.4 shows two disk-shaped molecules that have attracted much attention recently, benzene-hexa-*n*-alkanoate and hexa-*n*-alkoxy triphenylene.

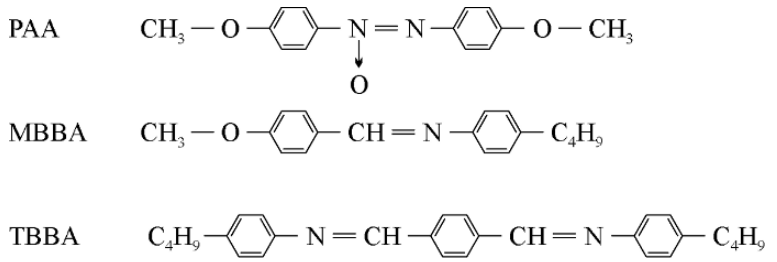


Fig. 2.3. Typical organic molecules exhibiting a liquid-crystalline phase

Phases can be classified according to the extent center-of-mass or orientational order is present in them. This classification goes back to G. Friedel (1922), although some liquid crystals have been known since the 19th century.[6]

[5] While liquid crystals made up of rod-like molecules were known for a long time, the first liquid crystal made up of disk-shaped molecules and exhibiting a new type of molecular arrangement was discovered by S. Chandrasekhar in 1977.

[6] The discovery of liquid crystals dates to 1888, when the Austrian botanist F. Reinitzer observed that cholesteryl benzoate had two distinct melting points: first a fuzzy liquid appeared, which suddenly became transparent at a higher temperature. One year later a German physicist, O. Lehmann demonstrated that this was a new phase of matter, possessing the characteristic properties of crystals and liquids alike, and he coined the term *liquid crystal*.

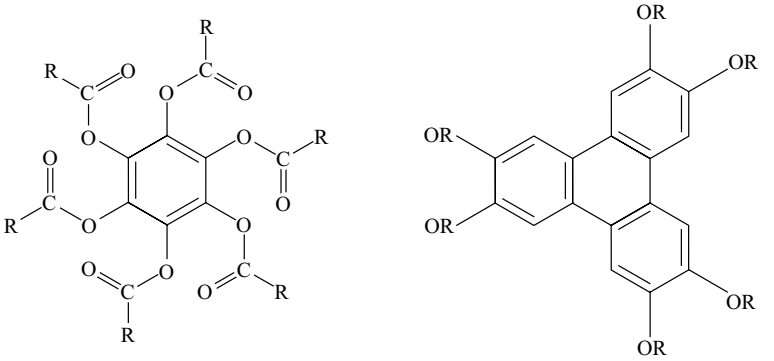


Fig. 2.4. Flat, disk-shaped molecules exhibiting a discotic liquid-crystalline phase. $R = \mathrm{C}_n\mathrm{H}_{2n+1}$ is an alkyl group

In *nematic*[7] (N) liquid crystals the positions of the centers of mass of rod-like molecules is disordered, while the orientation of their longitudinal axes is more or less ordered, as shown schematically in Fig. 2.5(a). Because of the disordered arrangement of the centers of mass, the system – just like any liquid – is invariant under arbitrary translations. Isotropy, i.e. invariance under rotations is broken: only rotations about the *director* – that is, the direction $\boldsymbol{n}$ of the long axis of the mesogenic groups –, and rotations through 180° about the directions perpendicular to it will take the system into itself. This is true even if the two tips of the molecules are inequivalent but half of the molecules point in one direction and the other half in the other. Thus, the director is not a true vector, since directions $\boldsymbol{n}$ and $-\boldsymbol{n}$ are equivalent. Therefore order cannot be characterized by the expectation value of the angle between the rod axes and a reference direction, rather the quantity

$$S = \tfrac{1}{2}\left\langle \left(3\cos^2\theta - 1\right)\right\rangle \tag{2.2.1}$$

plays the role of an order parameter. In the nematic phase the sample possesses cylindrical symmetry around the direction of the director. Because of the cylindrical symmetry – or the underlying preferential orientation of the molecules – the macroscopic properties of nematic liquid crystals show uniaxial anisotropy.

The nematic phase may appear in materials made up of flat, disk-shaped molecules as well. If the normal vectors to the disk planes are aligned but the arrangement of the centers of mass exhibits no order, as illustrated in Fig. 2.5(b), the system is said to be in a *discotic nematic* phase.[8]

[7] From the Greek words $\nu\eta\mu\alpha$ (nema) and $\nu\eta\mu\alpha\tau\iota\kappa\sigma\varsigma$ (nematikos), meaning thread, and woven, respectively.

[8] The nematic phase of rod-shaped molecules is sometimes called the *calamitic nematic phase* to emphasize the difference, although the calamitic and discotic nematic phases have the same symmetries, and the same notation N is recommended for both.

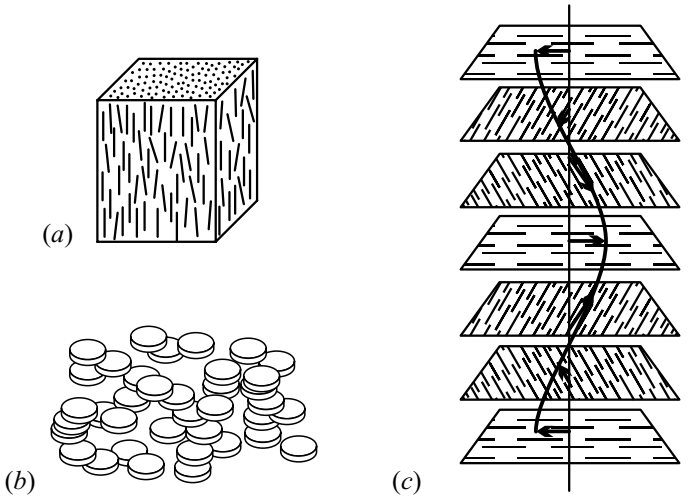


Fig. 2.5. Molecular arrangement in various liquid-crystalline phases: (*a*) nematic; (*b*) discotic nematic; (*c*) cholesteric phase

More recently nematic liquid crystals made up of rectangular, rather than rod- or disk-shaped molecules have been found. The arrangement of the centers of the rectangles is disordered but the molecules are packed together in such a way that their axes are more or less aligned. Cylindrical symmetry around the director is thus broken. This is called the *biaxial nematic* (N_b) phase.

The nematic phase is usually formed by molecules possessing left-right (up-down) symmetry. Chiral[9] – that is, not mirror symmetric – molecules can also form a nematic phase as long as left- and right-handed molecules are present in equal numbers – however, symmetries of the nematic phase are easily broken because of chirality, in which case another liquid-crystalline phase, the *cholesteric*[10] phase appears. In this phase the centers of mass are not ordered, and the director is not constant in space – as in the nematic phase – but rotates uniformly in a plane as one moves in the direction perpendicular to this plane. Choosing the z-axis along this direction,

$$\boldsymbol{n}(z) = [\hat{\boldsymbol{x}}\cos(qz+\phi),\ \hat{\boldsymbol{y}}\sin(qz+\phi),\ 0]\,. \tag{2.2.2}$$

The molecular arrangement is shown schematically in Fig. 2.5(*c*). In fact the molecules are not organized in planes; the change is continuous along the preferred axis. That is why the cholesteric phase is also called the *chiral nematic phase*, hence the notation N^*. As directions $\boldsymbol{n}$ and $-\boldsymbol{n}$ are equivalent,

[9] From the Greek word for hand, $\chi\epsilon\iota\rho$ (cheir), referring to the fact that hands are not mirror symmetric.

[10] The name refers to cholesterol esters, the first substances that were identified as liquid crystals.

the wavelength (chiral pitch) of the cholesteric phase is given by $\lambda = \pi/q$. This may depend sensitively on temperature but it is usually around 300–500 nm, that is, the pitch is on the same order as the wavelength of visible light. For this reason cholesteric liquid crystals exhibit unique optical properties. This is exploited in a number of applications, for example in liquid crystal displays (LCDs).

Smectic[11] liquid crystals come in several varieties. Figure 2.6 shows the structure of two of these, A and C (SmA and SmC, or S_A and S_C). Besides the alignment of the molecules, some kind of order is observed in the coordinates of the centers of mass, too, but only in one spatial direction. This means that molecules form layers. Within the layers, the arrangement of the molecular centers of mass is disordered, much in the same manner as in liquids, but their axes are along a preferred direction. The separation of adjacent layers is determined by the length of the molecules. The difference between smectic A and C phases is that the molecular axes are perpendicular to the plane of the layer in the former but not in the latter. In this case the average tilt angle in subsequent layers may be identical, reversed (giving a fishbone-like pattern of two neighboring layers), or it may even change periodically.

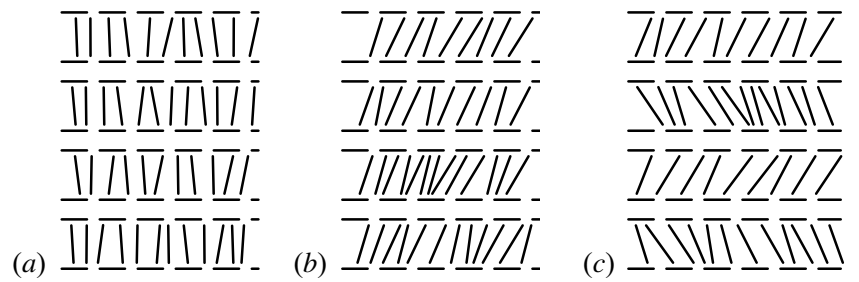


Fig. 2.6. Side view of the molecular arrangement in different smectic liquid crystals. (*a*): the smectic A phase; (*b*) and (*c*): two variants of the smectic C phase

Alternatively, one can say that smectic A and C liquid crystals are like solids in one direction but behave like liquids in the two perpendicular directions. Because of the incomplete order, the viscosity of liquid crystals is close to that of liquids. In materials possessing both smectic and nematic phases the former appears at a lower temperature, indicating that the smectic phase is more ordered than the nematic. When temperature is increased, a phase transition occurs: the layers disappear, and the symmetries of the nematic phase appear. Full translational symmetry is restored.

If individual layers are considered as rigid structural units, the smectic phase can be pictured as one in which these units are arranged regularly

[11] From the Greek words for soap, $\sigma\mu\eta\gamma\mu\alpha$ (smegma) and smeared, layered, $\sigma\mu\eta\kappa\tau o\varsigma$ (smektos).

along one direction. R. E. Peierls (1935) and L. D. Landau[12] (1937) showed that in one-dimensional systems short-range forces cannot give rise to such a long-range order: thermal fluctuations are so strong that they will always destroy the order. This is the so-called *Landau–Peierls instability*. Thus, strictly speaking, the sequence of smectic layers cannot show long-range order. A quasi-long-range order can, nevertheless exist. Experiments have revealed that along the direction of the layers the periodicity in the variations of the pair-correlation function is not perfect: the periodic function is multiplied by a function that decreases with a temperature-dependent power of distance. Moreover, the boundaries of smectic layers are smeared out – however, only to an extent that is much smaller than the width of individual layers.

There are further known variants of the smectic phase. However, their presentation would lead us too far afield – especially since some of them exhibit ordered crystalline structures in all three directions. We shall just mention one, the *smectic B* (SmB or S_B) or *hexatic phase*. In this phase, shown in Fig. 2.7(a), the arrangement of the molecules in a smectic layer is such that each molecule is surrounded by six neighbors. No hexagonal crystalline structure is formed, however, since the centers of mass exhibit only short- and not long-range order within the layers. On the other hand, the axes of the hexagons are aligned over long distances, so long-range orientational order is observed.

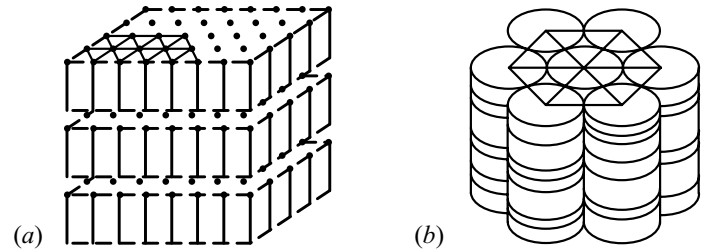


Fig. 2.7. Molecular arrangement in two liquid-crystalline phases: (a) smectic B phase; (b) discotic columnar phase

Hexagonal-type structures can also be formed by disk-shaped molecules. In the *discotic columnar phase* (or simply *columnar phase*) such flat molecules are stacked, forming a column. As shown in Fig. 2.7(b), these columns are packed locally hexagonally to ensure efficient space filling. In this hexagonal columnar phase (Col_h) the columns make up a regular hexagonal crystal. In two directions the substance behaves as a crystal, however, within the columns the order is only short-ranged, and thus typical liquid behavior is observed along this third direction. These columns do not always form hexagonal lattices:

[12] Lev Davidovich Landau was awarded the Nobel Prize in 1962, "for his pioneering theories for condensed matter, especially liquid helium".

other two-dimensional crystal structures are equally possible. For example, the $\mathrm{Col_r}$ phase features a rectangular lattice. Disk normals are not necessarily aligned with the column axis, either. This leads to an even greater wealth of possible structures.

A hierarchical order can be established among different liquid crystal types according to their symmetries. Starting out with the invariance under all rotations and translations in the liquid phase, symmetries are broken one by one, and so, through the symmetries characteristic of the nematic, smectic A, smectic C, and smectic B phases, one finally arrives at the symmetries of crystals. At this point all that is left is invariance under discrete translations and discrete rotations. This hierarchy exists not only in theory: it is also realized in some liquid crystals via subsequent phase transitions.

Orientationally Disordered Plastic Crystals

We have seen that in liquid crystals the arrangement of molecular centers of mass is more or less disordered, while molecular axes are aligned, and the system exhibits orientational order. Another kind of partially ordered state is possible in materials built up of large, almost spherical molecules that – due to their internal structure – are invariant only under discrete rotations. Here molecular centers of mass are arranged in a regular crystalline array while the orientation of molecular axes is disordered; center-of-mass order is thus not accompanied by orientational order. Because of their plastic properties such solids are called *plastic crystals*.

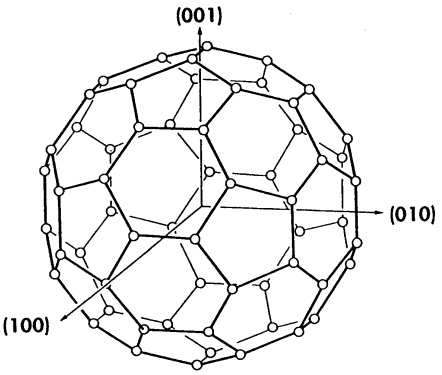


Fig. 2.8. The arrangement of carbon atoms in a C_{60} molecule

The C_{60} molecule shown in Fig. 2.8 has sixty carbon atoms arranged on the surface of a sphere in rings of five and six, much like a soccer ball or the geodesic domes of the architect R. Buckminster Fuller. From the latter's

name, the C_{60} molecule is also called a *fullerene* or a *buckyball*.[13] In the solid crystalline phase called *fullerite* C_{60} molecules are arranged in the vertices and face centers of the primitive cells of a cubic lattice. At room temperature molecular axes are not aligned along any particular direction relative to the axes of the lattice, moreover, they can rotate freely. Around 263 K a phase transition occurs and molecular axes become aligned. Instead of free rotation, the axis is allowed to rotate only around certain directions determined by symmetry. Only below this temperature is there a genuine crystalline order, including long-range orientational order.

Polymers constitute another interesting class of large organic molecules. Interactions between molecules will often get their chains intertwined, which is why arrangement into regular crystalline structures from the ordinary condensed phase requires special techniques. In general, only partial order is achieved. The interested reader is urged to consult the specialized literature.

Further Reading

1. P. M. Chaikin and T. C. Lubensky, *Principles of Condensed Matter Physics*, Cambridge University Press, Cambridge (2000).
2. S. Chandrasekhar, *Liquid Crystals*, Second Edition, Cambridge University Press, Cambridge (1992).
3. P. G. de Gennes and J. Prost, *The Physics of Liquid Crystals*, Second Edition, International Series of Monographs on Physics No. 83, Clarendon Press, Oxford (1995).
4. G. Strobl, *Condensed Matter Physics – Crystals, Liquids, Liquid Crystals, and Polymers*, Springer-Verlag, Berlin (2004).
5. B. K. Vainshtein, V. M. Fridkin, and V. L. Indenbom, *Modern Crystallography*, Volume 2; *Structure of Crystals*, Third edition, Springer-Verlag, Berlin (2000).

[13] Robert F. Curl Jr. (1933–), Sir Harold W. Kroto (1939–), and Richard E. Smalley (1943–2005) were awarded the 1996 Nobel Prize in Chemistry "for their discovery of fullerenes".

3

The Building Blocks of Solids

In what follows, of all condensed phases we shall study only the solid phase in the strictest sense of the word. Depending on what properties are examined, different theoretical methods and experimental techniques must be employed. The phenomenological approach is often sufficient for a theoretical description. This is especially true for mechanical properties, when the material is considered as an elastic continuum and the equations of elasticity can be used. Similarly, the conceptual framework and methods of thermodynamics or the electrodynamics of continuous media often provide a sufficiently good description of thermal, electric, or magnetic properties. Nevertheless to gain a deeper insight into the phenomena one has to study them more profoundly, starting off with the atomic structure of matter. In this theoretical approach solids have to be considered as highly complicated quantum mechanical systems comprised of many interacting particles.

Before tackling the many-particle problem, we shall separately summarize the behavior of ion cores and electrons that are not bound to atoms. In the present volume we shall deal almost exclusively with the role of ion cores. In this introductory chapter we shall first write down the Hamiltonian of the system of electrons and ions, and then recall, for future reference, some important results of quantum mechanics and atomic physics concerning the states of bound electrons in free atoms and ions.

3.1 Solids as Many-Particle Systems

The subject of theoretical investigations in solid-state physics is, in principle, very simple to present. Consider a quantum mechanical system that consists of N nuclei and N_e electrons. The Hamiltonian associated with the system is

denoted by $\mathcal{H}$. First, one has to determine the energy spectrum of the system, that is, the eigenvalues and eigenfunctions of the Schrödinger equation [1]

$$\mathcal{H}\Psi = E\Psi\,. \tag{3.1.1}$$

From these not only the ground-state characteristics can be read off: using the methods of statistical physics one can, in principle, determine all static and dynamic properties of the system in thermodynamic equilibrium at some finite temperature.

3.1.1 The Hamiltonian of Many-Particle Systems

The Hamiltonian can be split into three parts: nuclear and electronic contributions, and Coulomb interactions between the nuclei and the electrons:

$$\mathcal{H} = \mathcal{H}_{\text{nucl}} + \mathcal{H}_{\text{el}} + \mathcal{H}_{\text{el–nucl}}\,. \tag{3.1.2}$$

In the previous expression $\mathcal{H}_{\text{nucl}}$ contains the kinetic energies of the nuclei $i = 1, 2, \ldots, N$ of charge $Q_i e$ and mass M_i located at $\boldsymbol{R}_i$ plus their mutual Coulomb repulsion:

$$\mathcal{H}_{\text{nucl}} = \mathcal{H}_{\text{nucl}}^{\text{kin}} + \mathcal{H}_{\text{nucl–nucl}} = -\sum_{i=1}^{N} \frac{\hbar^2}{2M_i}\frac{\partial^2}{\partial \boldsymbol{R}_i^2} + \frac{1}{8\pi\epsilon_0}\sum_{\substack{i,j=1\\ i\neq j}}^{N} \frac{Q_i Q_j e^2}{|\boldsymbol{R}_i - \boldsymbol{R}_j|}\,. \tag{3.1.3}$$

The kinetic energies of and Coulomb interactions between electrons of charge $-e$ appear in a similar expression:

$$\mathcal{H}_{\text{el}} = \mathcal{H}_{\text{el}}^{\text{kin}} + \mathcal{H}_{\text{el–el}} = -\frac{\hbar^2}{2m_{\text{e}}}\sum_{i=1}^{N_{\text{e}}} \frac{\partial^2}{\partial \boldsymbol{r}_i^2} + \frac{1}{8\pi\epsilon_0}\sum_{\substack{i,j=1\\ i\neq j}}^{N_{\text{e}}} \frac{e^2}{|\boldsymbol{r}_i - \boldsymbol{r}_j|}\,, \tag{3.1.4}$$

while Coulomb interactions between the nuclei and the electrons are given by

$$\mathcal{H}_{\text{el–nucl}} = -\frac{1}{4\pi\epsilon_0}\sum_{i=1}^{N}\sum_{j=1}^{N_{\text{e}}} \frac{Q_i e^2}{|\boldsymbol{R}_i - \boldsymbol{r}_j|}\,. \tag{3.1.5}$$

We shall often encounter these expressions below. For simplicity, we introduce the notation

$$\tilde{e}^2 = \frac{e^2}{4\pi\epsilon_0}\,. \tag{3.1.6}$$

It will make many expressions appear as if they were written in Gaussian rather than rationalized units. For example, using this notation the electronic part of the Hamiltonian takes the form

[1] E. Schrödinger, 1926. Erwin Schrödinger (1887–1961) shared the Nobel Prize with Paul Adrien Maurice Dirac (1902–1984) in 1933 "for the discovery of new productive forms of atomic theory".

$$\mathcal{H}_{\mathrm{el}} = -\frac{\hbar^2}{2m_{\mathrm{e}}} \sum_{i=1}^{N_{\mathrm{e}}} \frac{\partial^2}{\partial \boldsymbol{r}_i^2} + \frac{1}{2} \sum_{\substack{i,j=1 \\ i \neq j}}^{N_{\mathrm{e}}} \frac{\tilde{e}^2}{|\boldsymbol{r}_i - \boldsymbol{r}_j|}\,, \tag{3.1.7}$$

which is commonly used in the literature even today.

Up to this point, the above Hamiltonian can equally well describe a gas, a plasma, a liquid, or a solid. In gases and liquids our task is much simpler as electrons do not move around freely but are bound to nuclei. Now interactions are relatively weak between neutral atoms or molecules, which can be considered as the new elementary units. Interactions continue to be weak even when the elementary units are multiply charged positive or negative ions, since the long-range mutual Coulomb forces between them are rapidly screened by other ions.

The situation is somewhat more complicated in plasmas. While a part of the electrons are again bound to nuclei, giving rise to relatively rigid ion cores, there are other electrons that are not bound and form an electron liquid. The description of the electronic states in solids is further complicated by the arbitrariness in choosing whether the electron belongs to the ion core or not. In general those electrons are considered to belong to the ion cores, which do not participate in chemical bonding.

In *ab initio* (first-principles) calculations the state of each electron has to be treated in the same manner. Less sophisticated calculations are considerably simplified by separating the contributions of ion cores and unbound electrons in the Hamiltonian. The Hamiltonian is then written as

$$\mathcal{H} = \mathcal{H}_{\mathrm{ion}} + \mathcal{H}_{\mathrm{el}} + \mathcal{H}_{\mathrm{el\text{–}ion}}\,, \tag{3.1.8}$$

where

$$\mathcal{H}_{\mathrm{ion}} = \mathcal{H}_{\mathrm{ion}}^{\mathrm{kin}} + \mathcal{H}_{\mathrm{ion\text{–}ion}} = -\sum_{i=1}^{N} \frac{\hbar^2}{2M_i} \frac{\partial^2}{\partial \boldsymbol{R}_i^2} + \frac{1}{2} \sum_{\substack{i,j=1 \\ i \neq j}}^{N} \frac{Z_i Z_j \tilde{e}^2}{|\boldsymbol{R}_i - \boldsymbol{R}_j|} \tag{3.1.9}$$

is the sum of the kinetic energies of ions of charge $Z_i e$ and mass M_i and their mutual Coulomb energies. The Hamiltonian of the system of electrons, $\mathcal{H}_{\mathrm{el}}$, can again be written in the form given in (3.1.4) and (3.1.7). However, summation is now not over all electrons but only over the N_{e} electrons that are not bound to ion cores. Finally, the Hamiltonian of the Coulomb interactions between ions and electrons is

$$\mathcal{H}_{\mathrm{el\text{–}ion}} = -\sum_{i=1}^{N} \sum_{j=1}^{N_{\mathrm{e}}} \frac{Z_i \tilde{e}^2}{|\boldsymbol{R}_i - \boldsymbol{r}_j|}\,. \tag{3.1.10}$$

Finding the solution to the eigenvalue problem of the total Hamiltonian is impossible for two reasons. First, the number of nuclei and electrons is excessively large. Second, the interaction is strong and long-ranged. This implies

that one must always employ some approximation method in the calculations; the (in)applicability of a particular method is often revealed only by the (dis)agreement with experiments.

The study of solids is considerably simplified if it is possible to assume that ions are essentially localized, even when their arrangement is not regular. Suppose the ionic positions are known; then the potential generated by them at point $\boldsymbol{r}$ is given by

$$U_{\text{ion}}(\boldsymbol{r}) = -\sum_{i=1}^{N} \frac{Z_i \tilde{e}^2}{|\boldsymbol{R}_i - \boldsymbol{r}|} \,. \tag{3.1.11}$$

Electrons move in this field. The great difficulty is that the mutual Coulomb interactions of the electrons cannot be written as a one-particle potential. Therefore approximation methods are necessary to study these interactions. As we proceed, we shall encounter more and more sophisticated methods.

3.1.2 Effects of Applied Fields

Solid-state physics is concerned not only with the properties of solids in thermal equilibrium: much more effort goes into the investigation of their behavior under external perturbations. In most cases this external perturbation is a temperature gradient or an applied electromagnetic field but it may also be an external pressure or a combination of them. The electromagnetic field can be a simple potential difference between two points of the sample, a uniform magnetic field, visible light, or a high-frequency electromagnetic radiation. Thus the Hamiltonian must very often be complemented by a term $\mathcal{H}_{\text{ext}}$ representing the applied field. As we shall see, an analysis of the response of the system to such external perturbations provides a deeper insight into its thermal equilibrium properties.

In the simplest case, when the sample is placed in an electromagnetic field and only the electrons' behavior is considered, an external potential appears in the Hamiltonian of the system of electrons. Besides, the kinetic energy term is modified: instead of the canonical momentum $\boldsymbol{p} = (\hbar/\mathrm{i})\boldsymbol{\nabla}$ it now contains the kinetic momentum $\boldsymbol{p} + e\boldsymbol{A}$.

To prove this consider the Newtonian equation of motion of a particle of mass m and charge q in external electric and magnetic fields,

$$m\frac{\mathrm{d}^2\boldsymbol{r}}{\mathrm{d}t^2} = q\big(\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B}\big) = q\big(\boldsymbol{E} + \dot{\boldsymbol{r}} \times \boldsymbol{B}\big) \,. \tag{3.1.12}$$

When the fields are expressed in terms of a scalar and a vector potential as

$$\boldsymbol{E} = -\operatorname{grad}\varphi(\boldsymbol{r}) - \frac{\partial \boldsymbol{A}}{\partial t} \,, \qquad \boldsymbol{B} = \operatorname{curl}\boldsymbol{A} \,, \tag{3.1.13}$$

it is readily seen that the above equation of motion can be derived in the Lagrangian formulation of mechanics from Lagrange's equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{\boldsymbol{r}}} - \frac{\partial \mathcal{L}}{\partial \boldsymbol{r}} = 0 \tag{3.1.14}$$

if the Lagrangian is chosen in the form

$$\mathcal{L} = \tfrac{1}{2}m\dot{\boldsymbol{r}}^2 - q\varphi(\boldsymbol{r}) + q\boldsymbol{v}\cdot\boldsymbol{A}\,, \tag{3.1.15}$$

where $\boldsymbol{v} = \dot{\boldsymbol{r}}$ is the velocity, and the relation

$$\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t} = \frac{\partial \boldsymbol{A}}{\partial t} + \frac{\partial \boldsymbol{A}}{\partial \boldsymbol{r}}\cdot\dot{\boldsymbol{r}} \tag{3.1.16}$$

is taken into account. Canonically conjugate momenta are customarily defined by the relation

$$\boldsymbol{p} = \frac{\partial \mathcal{L}}{\partial \dot{\boldsymbol{r}}} \tag{3.1.17}$$

in the Hamiltonian formulation. Thus, in the presence of an external electromagnetic field,

$$\boldsymbol{p} = m\dot{\boldsymbol{r}} + q\boldsymbol{A}\,. \tag{3.1.18}$$

When $\boldsymbol{r}$ and the conjugate momentum $\boldsymbol{p}$ are used as canonical variables, the Hamiltonian is obtained from

$$\mathcal{H} = \boldsymbol{p}\cdot\dot{\boldsymbol{r}} - \mathcal{L}\,. \tag{3.1.19}$$

In our case this leads to the expression

$$\mathcal{H} = \frac{1}{2m}\left(\boldsymbol{p} - q\boldsymbol{A}\right)^2 + q\varphi\,. \tag{3.1.20}$$

The quantum mechanical Hamiltonian has the same form, only the canonical commutation relations have to be required for $\boldsymbol{r}$ and $\boldsymbol{p}$. The quantity $\boldsymbol{p} - q\boldsymbol{A}$ is called the kinetic momentum, since according to (3.1.18) it is equal to $m\dot{\boldsymbol{r}}$, which appears in the kinetic energy formula. In contrast to the canonical momentum the kinetic momentum is gauge invariant, since the gauge transformation

$$\varphi \to \varphi' = \varphi - \frac{\partial f}{\partial t}\,, \qquad \boldsymbol{A} \to \boldsymbol{A}' = \boldsymbol{A} + \operatorname{grad} f \tag{3.1.21}$$

takes the wavefunction and an operator O into

$$\psi \to \psi' = U\psi\,, \qquad \mathcal{O} \to \mathcal{O}' = U\mathcal{O}U^{\dagger}\,, \tag{3.1.22}$$

where

$$U = \mathrm{e}^{\mathrm{i}qf/\hbar}\,. \tag{3.1.23}$$

When the potential due to the ions is combined with the external potential, and the notation $U(\boldsymbol{r})$ is introduced for the total potential, we obtain for electrons ($q = -e$)

$$\mathcal{H}_{\mathrm{el}} = \frac{1}{2m_{\mathrm{e}}} \sum_{i=1}^{N_{\mathrm{e}}} \left(\boldsymbol{p}_i + e\boldsymbol{A}(\boldsymbol{r}_i)\right)^2 + \sum_{i=1}^{N_{\mathrm{e}}} U(\boldsymbol{r}_i) + \frac{1}{2} \sum_{\substack{i,j=1 \\ i \neq j}}^{N_{\mathrm{e}}} \frac{\tilde{e}^2}{|\boldsymbol{r}_i - \boldsymbol{r}_j|} \,. \tag{3.1.24}$$

There is another term that has been ignored so far. Electrons possess an intrinsic angular momentum, a spin of one-half, and along with it a magnetic moment – thus they interact with the applied magnetic field. When the electron spin is treated as a dimensionless quantity (i.e., $\hbar$ is separated off),[2] the interaction with the magnetic field gives the following contribution to the Hamiltonian:

$$-g_{\mathrm{e}}\mu_{\mathrm{B}} \sum_{i=1}^{N_{\mathrm{e}}} \boldsymbol{B} \cdot \boldsymbol{s}_i \,, \tag{3.1.25}$$

where $g_{\mathrm{e}} \approx -2$ is the electron g factor, and $\boldsymbol{B}$ is the magnetic induction (also called the magnetic-flux density). The latter is related to the vector potential $\boldsymbol{A}$ by $\boldsymbol{B} = \operatorname{curl} \boldsymbol{A}$.

Electron spin plays an important role even in the absence of a magnetic field. To determine the wavefunction of a many-electron system, one has to take the Pauli exclusion principle[3] into account. This states that the wavefunction changes sign when the coordinates of any two electrons are interchanged. In addition to spatial coordinates, spin variables are also included among these coordinates. Thus even when the Hamiltonian is apparently spin-independent, the solution can depend on spin.

The magnetic field was considered as an external perturbation here. However, the field felt by the electrons may be partly due to the magnetic moment $\boldsymbol{\mu}_I$ of the nuclei. At the end of the chapter we shall present a derivation of the corresponding term in the Hamiltonian.

3.1.3 Relativistic Effects

When writing down the previous form of the Hamiltonian it was tacitly assumed that relativistic effects can be ignored. More precisely, it was taken into account that electrons have a spin – but apart from that the Schrödinger equation of nonrelativistic quantum mechanics was used. In solid-state physics this is often justified as the energy correction due to relativistic effects is four orders of magnitude smaller than the kinetic energy or the Coulomb energy.[4] However, these effects can be observed for electrons moving in the field of

[2] Throughout the book, we follow the convention that $\boldsymbol{s}$ and $\boldsymbol{l}$ are dimensionless quantities for the electron's intrinsic angular momentum (spin) and orbital angular momentum. Dimensionful quantities are obtained by multiplication with $\hbar$.

[3] W. Pauli, 1925. Wolfgang Pauli (1900–1958) was awarded the Nobel Prize in 1945 "for the discovery of the Exclusion Principle, also called the Pauli Principle".

[4] The energy correction is on the order of α^2, where $\alpha = \tilde{e}^2/\hbar c \approx 1/137$ is the fine-structure constant.

heavy ions. To demonstrate this, we shall consider a single electron, and start off with the Dirac equation.[5] When it is expanded in powers of $1/c$, the large component describing electrons can be separated from the small components describing positrons, thus an effective Hamiltonian can be obtained for electrons.

Keeping only first-order terms in $1/c$ we have

$$\mathcal{H} = m_\mathrm{e}c^2 + \frac{1}{2m_\mathrm{e}}\left(\boldsymbol{\sigma}\cdot(\boldsymbol{p}+e\boldsymbol{A})\right)^2 + U(\boldsymbol{r})\,, \tag{3.1.26}$$

where $\boldsymbol{\sigma}$ denotes the Pauli matrices given in (F.3.10). Making use of the identity

$$(\boldsymbol{\sigma}\cdot\boldsymbol{A})(\boldsymbol{\sigma}\cdot\boldsymbol{B}) = (\boldsymbol{A}\cdot\boldsymbol{B}) + \mathrm{i}\boldsymbol{\sigma}\cdot(\boldsymbol{A}\times\boldsymbol{B}) \tag{3.1.27}$$

among Pauli matrices, and neglecting the term $m_\mathrm{e}c^2$, we are left with

$$\mathcal{H}^{(1)} = \frac{1}{2m_\mathrm{e}}\big(\boldsymbol{p}+e\boldsymbol{A}(\boldsymbol{r})\big)^2 + U(\boldsymbol{r}) - g_\mathrm{e}\mu_\mathrm{B}\,\boldsymbol{B}\cdot\boldsymbol{s}\,, \tag{3.1.28}$$

the Pauli Hamiltonian of a single electron, since for spin-1/2 electrons $\boldsymbol{s} = \frac{1}{2}\boldsymbol{\sigma}$, $\mu_\mathrm{B} = e\hbar/2m_\mathrm{e}$, and for g_e the Dirac equation gives precisely -2.

Keeping terms of the next order in the power series of $1/c$ and taking into account the proper normalization of large and small components, two new terms appear,

$$\begin{aligned}\mathcal{H}^{(2)} = &-\frac{1}{8m_\mathrm{e}^3c^2}\big(\boldsymbol{\sigma}\cdot(\boldsymbol{p}+e\boldsymbol{A})\big)^4 \\ &-\frac{1}{8m_\mathrm{e}^2c^2}\Big[\big(\boldsymbol{\sigma}\cdot(\boldsymbol{p}+e\boldsymbol{A})\big),\big[\big(\boldsymbol{\sigma}\cdot(\boldsymbol{p}+e\boldsymbol{A})\big),U(\boldsymbol{r})\big]\Big]\,,\end{aligned} \tag{3.1.29}$$

where $[\,,\,]$ denotes the commutator. After some algebra, the following form is obtained:

$$\begin{aligned}\mathcal{H}^{(2)} = &-\frac{1}{8m_\mathrm{e}^3c^2}\big(\boldsymbol{\sigma}\cdot(\boldsymbol{p}+e\boldsymbol{A})\big)^4 + \frac{\hbar^2}{8m_\mathrm{e}^2c^2}\boldsymbol{\nabla}^2U(r) \\ &+\frac{\hbar}{4m_\mathrm{e}^2c^2}\boldsymbol{\sigma}\cdot\big(\boldsymbol{\nabla}U(\boldsymbol{r})\times(\boldsymbol{p}+e\boldsymbol{A})\big)\,.\end{aligned} \tag{3.1.30}$$

The first term, which can also be written as

$$-\frac{1}{2m_\mathrm{e}c^2}\left(\frac{(\boldsymbol{p}+e\boldsymbol{A})^2}{2m_\mathrm{e}} + \mu_\mathrm{B}\boldsymbol{\sigma}\cdot\boldsymbol{B}\right)^2\,, \tag{3.1.31}$$

contains the relativistic correction to the kinetic and magnetic energy. In solid-state physics this can be neglected. The second term is called the *Darwin term*;[6] in atomic Coulomb fields this is important only for s-electrons, since

[5] P. A. M. Dirac, 1928. See footnote on page 32.
[6] C. G. Darwin, 1928.

$\boldsymbol{\nabla}^2(1/r) \propto \delta(r)$. The most important is the third, spin-dependent term, which gives a correction that depends on the spin as well as the orbital motion of the electron. This is the *spin–orbit interaction term*.

The spin–orbit interaction term takes a particularly simple form when the potential $U(\boldsymbol{r})$ is central, i.e. depends only on $r = |\boldsymbol{r}|$, and so

$$\boldsymbol{\nabla} U(r) = \frac{\mathrm{d}U(r)}{\mathrm{d}r}\frac{\boldsymbol{r}}{r}\,. \tag{3.1.32}$$

If the vector potential is then neglected,

$$\mathcal{H}_{\text{s–o}} = \frac{\hbar}{4m_{\text{e}}^2c^2}\frac{1}{r}\frac{\mathrm{d}U(r)}{\mathrm{d}r}(\boldsymbol{r}\times\boldsymbol{p})\cdot\boldsymbol{\sigma} \tag{3.1.33}$$

is obtained. The term in parentheses is recognized as the orbital angular momentum operator of the electron, $\hbar\boldsymbol{l} = \boldsymbol{r}\times\boldsymbol{p}$. Now if $\boldsymbol{\sigma}$ is replaced by the electron spin operator, the expression takes the usual form

$$\mathcal{H}_{\text{s–o}} = \frac{\hbar^2}{2m_{\text{e}}^2c^2}\frac{1}{r}\frac{\mathrm{d}U(r)}{\mathrm{d}r}\,\boldsymbol{l}\cdot\boldsymbol{s}\,. \tag{3.1.34}$$

It can be shown that for an atom of atomic number Z the contribution of this term to the energy is proportional to Z^2, therefore it becomes important for the heavy elements at the high end of the periodic table. For lighter elements it can be considered as a weak perturbation to the interaction due to other electrons.

If $U(\boldsymbol{r})$ is the scalar potential of a uniform electric field $\boldsymbol{E}$, the state of the electron is determined by the Hamiltonian

$$\begin{aligned}\mathcal{H} = {} & \frac{1}{2m_{\text{e}}}\big[\boldsymbol{p}+e\boldsymbol{A}(\boldsymbol{r})\big]^2 + U(\boldsymbol{r}) - g_{\text{e}}\mu_{\text{B}}\boldsymbol{B}\cdot\boldsymbol{s} \\ & + \frac{\hbar e}{4m_{\text{e}}^2c^2}\boldsymbol{\sigma}\cdot\big[\boldsymbol{E}\times(\boldsymbol{p}+e\boldsymbol{A})\big]\,.\end{aligned} \tag{3.1.35}$$

When the electric field is directed along the z-axis and the magnetic field is absent, the spin–orbit interaction term is customarily written as

$$\mathcal{H}_{\text{s–o}} = \alpha\big(\sigma_x p_y - \sigma_y p_x\big)\,, \tag{3.1.36}$$

where the applied electric field is contained in the coefficient α. This term may play a major role in two-dimensional electron gases produced in semiconductor devices. In this context the term is called the *Rashba term*.[7]

3.2 The State of Ion Cores

The configuration of the electron shells in various (free) chemical elements is given in Appendix B. Data on the chemical reactivity of the elements have

[7] E. I. Rashba, 1960.

revealed that electrons in completely filled electron shells – whose state thus corresponds to that of the electrons in a noble-gas atom – belong to the ion core as they do not take part in chemical bonding. Things are different beyond the closed shells. Inside a solid, electrons in the not completely filled s- and p-shells can be easily stripped off the atom, leaving behind a charged ionized core. The situation is not so clear-cut for electrons in partially filled d- and f-shells. Sometimes they are thought to belong to the core, while in other cases they are considered to have broken free of the core, their state extending over the entire crystal. In the rest of this chapter we analyze the state of the ion core only, taking the number of core electrons from experiments.

It is well known that the energy spectrum can be exactly determined for a hydrogen atom with a single electron – whereas only an approximate calculation of the energy spectrum is possible for an ion core with Z electrons around the nucleus. In a zeroth-order approximation the many-electron atom can be viewed as if each electron were on a hydrogen-like orbit. (Restrictions of the Pauli exclusion principle apply.) The electron states obtained in this way are highly degenerate since their energy depends only on the principal quantum number n. When the Coulomb interaction between electrons is taken into account through an average central potential, the same energies will depend also on the azimuthal quantum number. Degenerate levels may split up, and then degeneracy will be partially lifted. In the next step fluctuations about the average potential and spin-orbit interaction are taken into account, which may give rise to further splitting, as energy will now depend on magnetic and spin quantum numbers as well. In many cases it is sufficient to determine this splitting in the first nonvanishing order in perturbation theory. However, some care must be taken, as it is not immaterial whether spin–orbit interactions are larger or smaller than the fluctuations about the average potential of the mutual Coulomb interactions among the electrons.

Apart from the heaviest elements, throughout the periodic table spin–orbit coupling is less important than fluctuations of the interelectronic Coulomb interaction. Thus in a first approximation spin–orbit interaction can be neglected, whereupon the orbital momenta $\boldsymbol{l}_i$ and spins $\boldsymbol{s}_i$ of individual electrons add up independently. The resultant angular momentum and spin[8] of the atom are then

$$\boldsymbol{L} = \sum_i \boldsymbol{l}_i \,, \qquad \boldsymbol{S} = \sum_i \boldsymbol{s}_i \,. \tag{3.2.1}$$

Since both $\boldsymbol{L}$ and $\boldsymbol{S}$ commute with the unperturbed Hamiltonian (i.e., the Hamiltonian without the spin–orbit interaction term), the electronic state is characterized by the quantum numbers L, S, m_L, and m_S derived from the eigenvalues of the operators L^2, S^2, L_z, and S_z. The states obtained this way are $(2L+1)(2S+1)$-fold degenerate. Next, spin–orbit interaction is taken into account. It is readily seen from the form

$$\boldsymbol{L} \cdot \boldsymbol{S} = \tfrac{1}{2}\big(L_+ S_- + L_- S_+\big) + L_z S_z \,, \tag{3.2.2}$$

[8] The rules for the addition of angular momenta are summarized in Appendix F.

where $L_\pm = L_x \pm \mathrm{i}L_y$ and $S_\pm = S_x \pm \mathrm{i}S_y$ that this interaction mixes states with different m_L and m_S, and splits multiply degenerate states into several levels. This gives the *fine structure* of electronic states. The energy shift of states with different values of the total angular momentum $\boldsymbol{J} = \boldsymbol{L} + \boldsymbol{S}$ is determined by the strength of the spin–orbit interaction. Then the state of the ion core is characterized by four quantum numbers: the magnitudes L, S, and J of the dimensionless vectors $\boldsymbol{L}$, $\boldsymbol{S}$, and $\boldsymbol{J}$ – more precisely, the eigenvalues of

$$\boldsymbol{L}^2 = L(L+1)\,, \qquad \boldsymbol{S}^2 = S(S+1)\,, \qquad \boldsymbol{J}^2 = J(J+1)\,, \tag{3.2.3}$$

plus the projection m_J of the component J_z along the quantization direction. Energy depends only on the quantum numbers L, S, and J, but not on m_J. These atomic terms – which are now only $2J+1$-fold degenerate for any given J – are customarily denoted by ${}^{2S+1}L_J$; however, for historical reasons, a letter rather than a number is used for L, according to the following code:

L	0	1	2	3	4	5	6
Letter	S	P	D	F	G	H	I

This coupling of the angular momenta of individual electrons is called the LS *coupling* or the *Russell–Saunders coupling*.[9]

Because of the strong spin–orbit coupling, this perturbative approach cannot be used for the characterization of the electronic state in elements at the high end of the periodic table. While it is possible to define a total orbital angular momentum $\boldsymbol{L}$ and a total spin $\boldsymbol{S}$, these are not meaningful quantum numbers as they do not commute with the Hamiltonian. In such cases one first has to add up the orbital angular momentum $\boldsymbol{l}_i$ and the spin $\boldsymbol{s}_i$ individually for each electron. These total angular momenta $\boldsymbol{j}_i = \boldsymbol{l}_i + \boldsymbol{s}_i$ commute with the $\boldsymbol{l}_i \cdot \boldsymbol{s}_i$ term of the Hamiltonian, and their sum gives the resultant angular momentum $\boldsymbol{J}$ of the system of electrons:

$$\boldsymbol{J} = \sum_i \boldsymbol{j}_i\,. \tag{3.2.4}$$

This is the so-called jj coupling. For simplicity, in the rest of this chapter we shall assume that LS coupling can be used.

In what follows, we shall first discuss how to determine the quantum numbers L and S of the total orbital angular momentum and spin of the ground-state configuration, and how $\boldsymbol{L}$ and $\boldsymbol{S}$ are added to obtain the ground-state quantum number J. Then we shall examine how core electrons are affected by an external magnetic field. Owing to their orbital motion and spin, electrons possess a magnetic moment as well, consequently ion cores show diamagnetic

[9] H. N. Russell and F. A. Saunders, 1925.

or paramagnetic behavior in an applied magnetic field. This phenomenon opens the way to the experimental determination of the electronic state of ion cores.

3.2.1 Hund's Rules

Consider an atom with some electrons on its outermost, partially filled shell. According to the rules of angular momentum addition, the sum of the orbital and spin angular momenta of individual electrons can lead to many different resultant $\boldsymbol{L}$s and $\boldsymbol{S}$es. Consider, for example, the configuration p^2 in which two p-electrons ($l = 1$) have the same principal quantum number. Denoting the states with magnetic quantum number $m_l = 1, 0, -1$ and spin quantum number $m_s = \pm\frac{1}{2}$ by $|m_l, m_s\rangle$, the two-electron state can be written as a linear combination of the states $|m_{l_1}, m_{s_1}\rangle|m_{l_2}, m_{s_2}\rangle$ – where, again, due attention must be paid to the antisymmetrization required by the Pauli exclusion principle. The six one-particle states will then give fifteen two-particle states, which are best grouped according to their total quantum numbers L and S. The total angular orbital momentum of two $l = 1$ particles is $L = 0,\ 1$, or 2. Similarly, the total spin quantum number of two spin-1/2 particles is either $S = 0$ or $S = 1$. The singlet state $S = 0$ is antisymmetric with respect to the interchange of the spin of the two particles. This can appear only in wavefunctions that are symmetric in their spatial variables, that is, the total orbital angular momentum must be $L = 0$ or $L = 2$. The triplet configuration $S = 1$ is symmetric in its spin variables, and thus the corresponding wavefunction must be antisymmetric in its spatial variables. This corresponds to an orbital angular momentum of $L = 1$. The allowed configurations and their multiplicities – $(2L + 1)(2S + 1)$ – are listed in Table 3.1.

Table 3.1. Allowed values of L and S and multiplicity in states with two p-electrons

L	S	Multiplicity	Notation
0	0	1	^{1}S
1	1	9	^{3}P
2	0	5	^{1}D

In the singlet state $S = 0$ the spatial parts of the wavefunctions of the two electrons can be identical, while in the $S = 1$ state they are necessarily different. The Coulomb repulsion between the electrons is expected to be weaker in this state, leading to a lower energy than in the singlet state.

Spin–orbit interaction causes further splitting of the ninefold degenerate energy level $L = 1$, $S = 1$. In the vector sum $\boldsymbol{J} = \boldsymbol{L} + \boldsymbol{S}$ the allowed values of J range from $|L - S|$ to $L + S$. Hence for a ^{3}P level the possible values of

J are 0, 1, and 2. The corresponding multiplicities – given by $J(J+1)$ – are 1, 3, and 5. Spin–orbit interaction thus splits the ninefold degenerate energy level into three levels, a nondegenerate, a threefold degenerate, and a fivefold degenerate level. Table 3.2 lists the possible configurations for two p-electrons when spin–orbit interaction is taken into account.

Table 3.2. Allowed values of L, S, and J, and multiplicity in states with two p-electrons

L	S	J	Multiplicity	Notation
0	0	0	1	$^1\mathrm{S}_0$
1	1	0	1	$^3\mathrm{P}_0$
1	1	1	3	$^3\mathrm{P}_1$
1	1	2	5	$^3\mathrm{P}_2$
2	0	2	5	$^1\mathrm{D}_2$

Experience shows that of the five allowed states listed above, $^3\mathrm{P}_0$ (with quantum numbers ($L=1$, $S=1$, $J=0$) is the ground state. Such experimental findings are generalized by *Hund's rules*,[10] established in the golden age of spectroscopy, which provide a set of rules to determine the quantum numbers of the lowest-energy configuration.

1.) To minimize the Coulomb repulsion due to the overlap of their wavefunctions, electrons tend to occupy states with different magnetic quantum numbers m_l (as long as this is possible). As we shall see in Chapter 4, even though the Coulomb interaction is spin independent, the indistinguishability of quantum mechanical particles nevertheless gives rise to a spin-dependent interaction, the *exchange interaction*. As the exchange integral of two orthogonal states is positive, in less than half filled shells it is energetically more favorable to have the spins of all the electrons aligned, which of course requires that their magnetic quantum numbers m_l be different. When a shell is more than half filled, there will certainly be electrons with identical m_l but opposite m_s quantum numbers. Nevertheless even in this case the lowest-energy state is found to be the one in which the total spin has the largest possible value. If n of the $2(2l+1)$ possible states of the shell with azimuthal quantum number l are filled, the quantum number S of the ground state is

$$S = \tfrac{1}{2}\big[2l+1-|2l+1-n|\big] = \begin{cases} \frac{1}{2}n & n \le 2l+1\,, \\ 2l+1-\frac{1}{2}n & n \ge 2l+1\,. \end{cases} \tag{3.2.5}$$

2.) Once the total S is determined, the total orbital angular momentum L has to be chosen as large as possible. If there are n electrons on the shell

[10] F. Hund, 1925.

with azimuthal quantum number l, the total orbital angular momentum L is given by

$$L = S|2l+1-n| = \begin{cases} \frac{1}{2}n(2l+1-n) & n \le 2l+1\,, \\ \left(2l+1-\frac{1}{2}n\right)(n-2l-1) & n \ge 2l+1\,. \end{cases} \tag{3.2.6}$$

This arrangement is also dictated by the Coulomb energy – although it cannot be illustrated with an intuitive picture.

3.) Owing to the spin–orbit interaction, $\boldsymbol{L}$ and $\boldsymbol{S}$ are no longer conserved, only their sum, the total angular momentum $\boldsymbol{J} = \boldsymbol{L} + \boldsymbol{S}$ is. For fixed values of S and L the allowed values of the quantum number J (i.e., the length of $\boldsymbol{J}$) range from $|L-S|$ to $L+S$. The multiplicities of the split levels indeed add up to $(2L+1)(2S+1)$, the multiplicity prior to splitting:

$$\sum_{J=|L-S|}^{L+S} J(J+1) = (2L+1)(2S+1)\,. \tag{3.2.7}$$

The sign of the coupling constant of the spin–orbit interaction determines which of these has a lower energy. It is possible to provide theoretical justification to the empirical rule asserting that if a shell of azimuthal quantum number l is less than half filled (that is, when the number of electrons is smaller than $2l+1$) then the lowest-energy state has $J = |L-S|$, while if it is more then half filled then it has $J = L+S$. In case of n electrons

$$J = S|2l-n| = \begin{cases} \frac{1}{2}n(2l-n) & n \le 2l\,, \\ \left(2l+1-\frac{1}{2}n\right)(n-2l) & n > 2l\,. \end{cases} \tag{3.2.8}$$

The quantum numbers S, L, and J of the lowest-energy state satisfying Hund's rules are given in Table 3.3 for the case when the 10 states of the d-shell ($l = 2$) are gradually filled in.

It should be kept in mind that the table shows only schematically how the quantum numbers S and L could be determined for a given filling of the level. The wavefunction itself cannot be inferred directly from it. For this one has to use the angular momentum addition rules given in Appendix F. E.g., if there is a single electron on the 3d level, the wavefunctions of the fourfold degenerate ${}^2\mathrm{D}_{3/2}$ multiplet are linear combinations of the ten $|m_l, m_s\rangle$ states with orbital and spin quantum numbers $l = 2$ and $s = 1/2$. The state with maximal $m_J = 3/2$ is

$$|J = 3/2, m_J = 3/2\rangle = \frac{2}{\sqrt{5}}|2, -1/2\rangle - \frac{1}{\sqrt{5}}|1, 1/2\rangle\,. \tag{3.2.9}$$

The three other states that correspond to $m_J = 1/2, -1/2$, and $-3/2$ can be obtained from this state by successive applications of the lowering operator $J^- = L^- + S^-$.

Table 3.3. The n-electron ground states of the d-shell, as required by Hund's rules

n	2	1	0	−1	−2	S	L	J	$^{2S+1}L_J$
1	↑					1/2	2	3/2	$^2\mathrm{D}_{3/2}$
2	↑	↑				1	3	2	$^3\mathrm{F}_2$
3	↑	↑	↑			3/2	3	3/2	$^4\mathrm{F}_{3/2}$
4	↑	↑	↑	↑		2	2	0	$^5\mathrm{D}_0$
5	↑	↑	↑	↑	↑	5/2	0	5/2	$^6\mathrm{S}_{5/2}$
6	↑↓	↑	↑	↑	↑	2	2	4	$^5\mathrm{D}_4$
7	↑↓	↑↓	↑	↑	↑	3/2	3	9/2	$^4\mathrm{F}_{9/2}$
8	↑↓	↑↓	↑↓	↑	↑	1	3	4	$^3\mathrm{F}_4$
9	↑↓	↑↓	↑↓	↑↓	↑	1/2	2	5/2	$^2\mathrm{D}_{5/2}$
10	↑↓	↑↓	↑↓	↑↓	↑↓	0	0	0	$^1\mathrm{S}_0$

The energy splitting between levels with different quantum numbers L and S is typically on the order of an electronvolt, since it is due to the Coulomb interaction and its consequence, the exchange interaction. However, the splitting between states with different Js is at most 0.1 eV, since this is the result of the spin–orbit interaction. In spectroscopy, energy differences are customarily given by the wave number (in cm^{-1} units) of the photon associated with the transition. According to the conversion formulas presented in Appendix A, $1\,\mathrm{eV} \approx 8 \times 10^3\,\mathrm{cm}^{-1}$. Thus, the typical energy difference between multiplets with different L and S quantum numbers is 10^3–$10^4\,\mathrm{cm}^{-1}$, while between states with different Js it is usually 10^2–$10^3\,\mathrm{cm}^{-1}$.

As thermal energy reaches 0.1 eV at temperatures around 10^3 K, higher multiplets are usually not excited thermally because of the large energy difference. Therefore in the overwhelming majority of cases it is sufficient to consider only levels whose quantum numbers S and L are determined by Hund's first and second rules. Moreover, it is often enough to consider only the ground-state multiplet whose quantum number J is determined from Hund's third rule.

3.2.2 Angular Momentum and Magnetic Moment

It is known from the Zeeman[11] splitting of energy levels that the angular momentum $\hbar\boldsymbol{L}$ of an ion is always accompanied by a magnetic moment

$$\boldsymbol{\mu} = -\mu_\mathrm{B}\boldsymbol{L}\,, \tag{3.2.10}$$

where $\mu_\mathrm{B} = e\hbar/(2m_\mathrm{e})$ is the Bohr magneton. We shall shortly see how this is derived from the Schrödinger equation of an electron in a magnetic field. Using the physical angular momentum $\hbar\boldsymbol{L}$ instead of the dimensionless $\boldsymbol{L}$, the previous formula is alternatively written as

[11] See footnote on page 2.

$$\boldsymbol{\mu} = -\gamma_{\mathrm{e}} \hbar \boldsymbol{L} \,, \tag{3.2.11}$$

where γ_{e} is the gyromagnetic (magnetomechanical) ratio of the electron. Its value is

$$\gamma_{\mathrm{e}} = \frac{e}{2m_{\mathrm{e}}} = 1.760\,860 \times 10^{11}\,\mathrm{s}^{-1}\,\mathrm{T}^{-1} \,. \tag{3.2.12}$$

Owing to its orbital motion, an electron in a state characterized by the quantum numbers n, l, and m_l has a magnetic moment of magnitude

$$\mu_l = \mu_{\mathrm{B}} \sqrt{l(l+1)} \,, \tag{3.2.13}$$

while the z component of this magnetic moment is

$$\mu_l^z = -\mu_{\mathrm{B}} m_l \,. \tag{3.2.14}$$

The electron's intrinsic angular momentum, its spin $\boldsymbol{s}$ is accompanied by an intrinsic magnetic moment $\boldsymbol{\mu}_{\mathrm{s}}$. Just like orbital angular momentum and magnetic moment, spin and intrinsic magnetic moment are also oppositely directed. However, as the *Einstein–de Haas experiment*[12] revealed, the gyromagnetic ratio is now different from that obtained for the orbital angular momentum. The same conclusion is drawn from the Dirac equation. As mentioned earlier, the following relation holds between the dimensionless spin and the intrinsic magnetic moment in the nonrelativistic limit:

$$\boldsymbol{\mu}_{\mathrm{s}} = -2\mu_{\mathrm{B}} \boldsymbol{s} \,. \tag{3.2.15}$$

This relation is slightly modified by the radiation corrections of quantum electrodynamics:

$$\boldsymbol{\mu}_{\mathrm{s}} = g_{\mathrm{e}} \mu_{\mathrm{B}} \boldsymbol{s} \,, \tag{3.2.16}$$

where calculations and measurements give the same value for g_{e}, the electron g-factor:

$$g_{\mathrm{e}} = -2 \left[1 + \frac{\alpha}{2\pi} + \mathcal{O}(\alpha^2) \right] = -2.002\,319 \,. \tag{3.2.17}$$

As before, α is the fine-structure constant. To a very good approximation, g_{e} is equal to -2.[13]

Thus the z component of the magnetic moment in a state with spin quantum number $m_s = \pm\frac{1}{2}$ takes the value

$$\mu_{\mathrm{s}}^z = \pm\tfrac{1}{2} g_{\mathrm{e}} \mu_{\mathrm{B}} \approx \mp \mu_{\mathrm{B}} \,. \tag{3.2.18}$$

[12] A. Einstein and J. W. de Haas, 1915. Albert Einstein (1879–1955) was awarded the Nobel Prize in 1921, "for his services to Theoretical Physics, and especially for his discovery of the law of the photoelectric effect".

[13] Note that while in its rigorous definition g_{e} is negative, most solid-state physics references nevertheless use its absolute value. For easy comparison, we shall often use the notation $|g_{\mathrm{e}}|$ below.

3.2.3 The Magnetic Hamiltonian of Atomic Electrons

To determine the magnetic properties of atomic electrons, consider an electron moving in a potential $U(\boldsymbol{r})$ and an applied magnetic field $\boldsymbol{B}(\boldsymbol{r})$. If the latter is represented by a vector potential $\boldsymbol{A}(\boldsymbol{r})$, the canonical momentum $\boldsymbol{p}$ is replaced by the kinetic momentum $\boldsymbol{p} + e\boldsymbol{A}$ in the electron's Hamiltonian. When the interaction between the magnetic field and the magnetic moment due to the electron spin is taken into account, the behavior of a single electron is governed by the Hamiltonian (3.1.28),

$$\mathcal{H} = \frac{1}{2m_\mathrm{e}}\big(\boldsymbol{p} + e\boldsymbol{A}(\boldsymbol{r})\big)^2 + U(\boldsymbol{r}) + |g_\mathrm{e}|\mu_\mathrm{B}\boldsymbol{B}\cdot\boldsymbol{s}\,. \tag{3.2.19}$$

When expanding the quadratic expression of the kinetic energy, it should be remembered that the order of the operators $\boldsymbol{p}$ and $\boldsymbol{A}$ is not immaterial. It can be seen from the form of the momentum operator,

$$\boldsymbol{p} = \frac{\hbar}{\mathrm{i}}\boldsymbol{\nabla} = \frac{\hbar}{\mathrm{i}}\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) \tag{3.2.20}$$

that $\boldsymbol{p}$ and $\boldsymbol{A}$ do not commute in general:

$$[\boldsymbol{p}, \boldsymbol{A}] = \frac{\hbar}{\mathrm{i}}\,\mathrm{div}\,\boldsymbol{A}\,. \tag{3.2.21}$$

To simplify calculations it is useful to choose the *Coulomb gauge* (also known as *radiation gauge* or *transverse gauge*) $\mathrm{div}\,\boldsymbol{A} = 0$. When the magnetic field is uniform, one possible choice for the vector potential is

$$\boldsymbol{A} = \tfrac{1}{2}\,[\boldsymbol{B}\times\boldsymbol{r}]\,. \tag{3.2.22}$$

This is the so-called *symmetric gauge*. It is easily seen that both $\boldsymbol{B} = \mathrm{curl}\,\boldsymbol{A}$ and the condition $\mathrm{div}\,\boldsymbol{A} = 0$ are satisfied.

Substituting this form of the vector potential into the first term of the Hamiltonian (3.2.19), the following form is obtained:

$$\begin{aligned}\frac{1}{2m_\mathrm{e}}\left(\boldsymbol{p} - \frac{e}{2}[\boldsymbol{r}\times\boldsymbol{B}]\right)^2 &= \frac{\boldsymbol{p}^2}{2m_\mathrm{e}} - \frac{e}{2m_\mathrm{e}}\boldsymbol{p}\cdot[\boldsymbol{r}\times\boldsymbol{B}] + \frac{e^2}{8m_\mathrm{e}}\,[\boldsymbol{r}\times\boldsymbol{B}]^2 \\ &= \frac{\boldsymbol{p}^2}{2m_\mathrm{e}} + \frac{e}{2m_\mathrm{e}}\,[\boldsymbol{r}\times\boldsymbol{p}]\cdot\boldsymbol{B} + \frac{e^2}{8m_\mathrm{e}}\,[\boldsymbol{r}\times\boldsymbol{B}]^2\,.\end{aligned} \tag{3.2.23}$$

The angular momentum $\boldsymbol{r}\times\boldsymbol{p} = \hbar\boldsymbol{l}$ is immediately recognized in the middle term, and the Bohr magneton in its coefficient. The total Hamiltonian can therefore be written as

$$\begin{aligned}\mathcal{H} &= \frac{\boldsymbol{p}^2}{2m_\mathrm{e}} + \mu_\mathrm{B}\boldsymbol{l}\cdot\boldsymbol{B} + \frac{e^2}{8m_\mathrm{e}}\,[\boldsymbol{r}\times\boldsymbol{B}]^2 + U(\boldsymbol{r}) + |g_\mathrm{e}|\mu_\mathrm{B}\boldsymbol{B}\cdot\boldsymbol{s} \\ &= \frac{\boldsymbol{p}^2}{2m_\mathrm{e}} + U(\boldsymbol{r}) + \mu_\mathrm{B}(\boldsymbol{l} + |g_\mathrm{e}|\boldsymbol{s})\cdot\boldsymbol{B} + \frac{e^2}{8m_\mathrm{e}}\,[\boldsymbol{r}\times\boldsymbol{B}]^2\,.\end{aligned} \tag{3.2.24}$$

Using $\boldsymbol{r}_\perp$, the component of the electron's position vector perpendicular to the magnetic field, the vector product in the last term can be rewritten as $|\boldsymbol{r} \times \boldsymbol{B}| = |\boldsymbol{B}| \cdot |\boldsymbol{r}_\perp|$, leading to

$$\mathcal{H} = \frac{\boldsymbol{p}^2}{2m_\mathrm{e}} + U(\boldsymbol{r}) + \mu_\mathrm{B}(\boldsymbol{l} + |g_\mathrm{e}|\boldsymbol{s}) \cdot \boldsymbol{B} + \frac{e^2}{8m_\mathrm{e}} B^2 \boldsymbol{r}_\perp^2 \,. \tag{3.2.25}$$

This expression is readily generalized for the case when the atom has Z electrons. Labeling the electrons by i, one starts off with the counterpart of the Hamilton (3.2.19),

$$\mathcal{H} = \sum_{i=1}^{Z} \left[\frac{1}{2m_\mathrm{e}} \left(\boldsymbol{p}_i + e\boldsymbol{A}(\boldsymbol{r}_i) \right)^2 + U(\boldsymbol{r}_i) + |g_\mathrm{e}| \mu_\mathrm{B} \boldsymbol{B} \cdot \boldsymbol{s}_i \right] . \tag{3.2.26}$$

Following the same steps as above, the final result is now

$$\mathcal{H} = \sum_{i=1}^{Z} \frac{\boldsymbol{p}_i^2}{2m_\mathrm{e}} + \mu_\mathrm{B} \left(\boldsymbol{L} + |g_\mathrm{e}| \boldsymbol{S} \right) \cdot \boldsymbol{B} + \frac{e^2}{8m_\mathrm{e}} B^2 \sum_{i=1}^{Z} \boldsymbol{r}_{i\perp}^2 \,, \tag{3.2.27}$$

where

$$\hbar \boldsymbol{L} = \sum_{i=1}^{Z} \hbar \boldsymbol{l}_i = \sum_{i=1}^{Z} \left(\boldsymbol{r}_i \times \boldsymbol{p}_i \right) \tag{3.2.28}$$

is the total orbital angular momentum due to the orbital motion of the electrons, and

$$\boldsymbol{S} = \sum_{i=1}^{Z} \boldsymbol{s}_i \tag{3.2.29}$$

is the total spin of the system of electrons.

3.2.4 Magnetization and Susceptibility

The Hamiltonian of the electrons contains terms that are linear and quadratic in the magnetic field. The linear term is present when the total magnetic moment $-\mu_\mathrm{B}(\boldsymbol{L} + |g_\mathrm{e}|\boldsymbol{S})$ is finite (nonzero). In this case this moment tends to be aligned with the field. This *paramagnetic* contribution is dominant unless the total moment vanishes for some reason. Before analyzing the conditions for this, we have to say a few words about the determination of magnetization and susceptibility.

If the energies E_i of the states of the system are shifted under the influence of a magnetic field $\boldsymbol{B}$, the free energy also depends on the field. Using the partition function

$$Z = \sum_i \mathrm{e}^{-E_i(\boldsymbol{B})/k_\mathrm{B}T} \,, \tag{3.2.30}$$

the Helmholtz free energy can be written as

$$F = -k_{\mathrm{B}}T \ln Z = -k_{\mathrm{B}}T \ln \left\{ \sum_i \mathrm{e}^{-E_i(\boldsymbol{B})/k_{\mathrm{B}}T} \right\} . \tag{3.2.31}$$

Magnetization is the negative partial derivative of the free energy density $f = F/V$ with respect to magnetic induction:

$$\boldsymbol{M}(\boldsymbol{B}, T) = -\frac{\partial f}{\partial \boldsymbol{B}} = -\frac{1}{V}\frac{\partial F}{\partial \boldsymbol{B}} . \tag{3.2.32}$$

This can equally be written as

$$\boldsymbol{M}(\boldsymbol{B}, T) = \frac{\sum_i \boldsymbol{M}_i(\boldsymbol{B}) \mathrm{e}^{-E_i(\boldsymbol{B})/k_{\mathrm{B}}T}}{\sum_i \mathrm{e}^{-E_i(\boldsymbol{B})/k_{\mathrm{B}}T}} , \tag{3.2.33}$$

where

$$\boldsymbol{M}_i(\boldsymbol{B}) = -\frac{1}{V}\frac{\partial E_i(\boldsymbol{B})}{\partial \boldsymbol{B}} \tag{3.2.34}$$

is the contribution of the ith state to magnetization.

The customary definition of magnetic susceptibility – the response of the system to a magnetic field – is not based on the straightforward choice

$$\boldsymbol{\chi}' = \left. \frac{\partial \boldsymbol{M}}{\partial \boldsymbol{B}} \right|_{\boldsymbol{B}=0} \tag{3.2.35}$$

but rather on

$$\boldsymbol{\chi} = \left. \frac{\partial \boldsymbol{M}}{\partial \boldsymbol{H}} \right|_{\boldsymbol{H}=0} . \tag{3.2.36}$$

The derivative can also be taken at finite values of the field, leading to a field-dependent susceptibility. Susceptibility, with its ordinary definition, is the weak-field limit of the latter. In certain systems nonlinear contributions to the magnetization are also used to characterize magnetic properties.

Susceptibility defined in this manner is usually a tensor quantity,

$$\chi_{\alpha\beta} = \left. \frac{\partial M_\alpha}{\partial H_\beta} \right|_{\boldsymbol{H}=0} . \tag{3.2.37}$$

However, in the isotropic case off-diagonal elements vanish and diagonal ones are identical – so susceptibility can be considered a scalar. The dimensionless scalar susceptibility is related to magnetic permeability via

$$\mu = \mu_{\mathrm{r}}\mu_0 = (1 + \chi)\mu_0 . \tag{3.2.38}$$

Apart from some magnetically ordered materials (to be discussed in Chapter 14), the magnetization of solids is usually rather small ($\chi \ll 1$). As χ' of (3.2.26) is much more easy to determine theoretically, the relation $\boldsymbol{B} \approx \mu_0 \boldsymbol{H}$ permits one to obtain the dimensionless susceptibility χ from

$$\chi = \mu_0 \chi' . \tag{3.2.39}$$

When this approximation is justified, susceptibility can be calculated from the free-energy expression as

$$\chi = -\frac{\mu_0}{V}\frac{\partial^2 F(B)}{\partial B^2}\,. \tag{3.2.40}$$

As we started off with the free-energy density, χ above is frequently referred to as the *volume susceptibility* or *bulk susceptibility*.

Note that instead of the dimensionless susceptibility often χ_{mass}, the *mass* or *specific susceptibility*, or χ_{mol}, the *molar susceptibility* is given as experimental data. The former can be obtained by dividing χ by the density, $\chi_{\mathrm{mass}} = \chi/\rho$, while the latter by multiplying χ by the molar volume, $\chi_{\mathrm{mol}} = \chi V_{\mathrm{mol}}$.

3.2.5 Langevin or Larmor Diamagnetism

If the outermost electron shell of an atom or ion is completely filled, as in noble gases (He, Ne, Ar, etc.) or ionic crystals (Na^+, Cl^-, etc.), the ground-state quantum numbers are

$$L = S = J = 0\,. \tag{3.2.41}$$

When the energy correction due to the magnetic field is calculated to the lowest order in perturbation theory, the second term in (3.2.27) does not contribute. The only surviving term is proportional to B^2. For a single atom with Z electrons, the change in energy is

$$\Delta E = \frac{e^2}{8m_{\mathrm{e}}}B^2\Big\langle 0\Big|\sum_{i=1}^{Z}\boldsymbol{r}_{i\perp}^2\Big|0\Big\rangle\,. \tag{3.2.42}$$

For completely filled electron shells the electron distribution is spherically symmetric. It is an elementary exercise to show that in such cases the mean square of the position vector component perpendicular to the magnetic field is equal to two-thirds of the mean square of the magnitude of the position vector. Thus,

$$\Delta E = \frac{e^2}{12m_{\mathrm{e}}}B^2\Big\langle 0\Big|\sum_{i=1}^{Z}\boldsymbol{r}_{i}^2\Big|0\Big\rangle\,. \tag{3.2.43}$$

In systems containing N atoms in volume V, this shift of the ground-state energy gives rise to a susceptibility on the order of

$$\boxed{\chi = -\mu_0\frac{N}{V}\frac{\partial^2\Delta E}{\partial B^2} = -\frac{e^2\mu_0}{6m_{\mathrm{e}}}\frac{N}{V}\Big\langle 0\Big|\sum_{i=1}^{Z}\boldsymbol{r}_{i}^2\Big|0\Big\rangle\,,} \tag{3.2.44}$$

as at $T = 0$ susceptibility should be calculated from the energy instead of the free energy. As the negative sign shows, ions with closed shells behave diamagnetically. This is the *Langevin* or *Larmor diamagnetism*. The latter name

reflects that the phenomenon lends itself to classical interpretation based on Larmor's theorem on the motion of electrons in a magnetic field[14] – however, the interpretation is due to P. Langevin (1905). This asserts that when an electron moving in a closed orbit around a nucleus is placed in a magnetic field, the angular frequency of the orbiting electron will be shifted by the Larmor frequency $\omega_{\mathrm{L}} = eB/2m_{\mathrm{e}}$, inducing a magnetic moment that is proportional to the applied magnetic field. According to Lenz's law, the induced moment will oppose the applied field. This is the root of diamagnetism.

To obtain an order-of-magnitude estimate for the diamagnetic energy correction, we shall approximate the radius of atomic orbits by the Bohr radius $a_0 = 4\pi\epsilon_0\hbar^2/m_{\mathrm{e}}e^2 = \hbar^2/m_{\mathrm{e}}\tilde{e}^2$. The energy shift per electron is then

$$\Delta E_{\mathrm{dia}} \approx \frac{e^2}{12m_{\mathrm{e}}} B^2 a_0^2 \,. \tag{3.2.45}$$

In terms of the Larmor frequency,

$$\Delta E_{\mathrm{dia}} \approx \frac{e^2}{12m_{\mathrm{e}}} B^2 a_0 \frac{\hbar^2}{m_{\mathrm{e}}\tilde{e}^2} = \frac{1}{3}\hbar^2 \left(\frac{eB}{2m_{\mathrm{e}}}\right)^2 \frac{a_0}{\tilde{e}^2} = \frac{1}{3}\left(\hbar\omega_{\mathrm{L}}\right)^2 \left(\frac{\tilde{e}^2}{a_0}\right)^{-1} . \tag{3.2.46}$$

While $\tilde{e}^2/a_0$ is on the order of atomic energies, the magnetic energy $\hbar\omega_{\mathrm{L}}$ is much smaller. Even for fields as strong as $B \sim 1\,\mathrm{tesla}$ the difference is five orders of magnitude. In SI units, the dimensionless susceptibility χ is on the order of 10^{-9} for noble gases. This corresponds to a molar susceptibility of 10^{-10}–$10^{-11}\,\mathrm{m}^3/\mathrm{mol}$, and a specific susceptibility on the order of $10^{-9}\,\mathrm{m}^3/\mathrm{kg}$. Molar and specific susceptibilities for noble gases are listed in Table 3.4. When working with CGS units, the numerical value of the susceptibility differs by a factor of 4π: $\chi_{\mathrm{CGS}} = \chi_{\mathrm{SI}}/4\pi$.

Table 3.4. Molar and specific susceptibilies of noble gases at room temperature

	χ_{mol} [SI] $(10^{-12}\,\mathrm{m}^3/\mathrm{mol})$	χ_{mol} [CGS] $(10^{-6}\,\mathrm{cm}^3/\mathrm{mol})$	χ_{mass} [SI] $(10^{-9}\,\mathrm{m}^3/\mathrm{kg})$
He	−25.4	−2.02	−5.9
Ne	−87.5	−6.96	−4.2
Ar	−243	−19.3	−6.16
Kr	−364	−29.0	−4.32
Xe	−572	−45.5	−4.20

In ionic crystals, the dimensionless susceptibility is on the order of 10^{-4} to 10^{-6} – e.g., for NaCl $\chi = -13.9 \times 10^{-6}$, which corresponds to a molar susceptibility of $\chi_{\mathrm{mol}} = -3.75 \times 10^{-10}\,\mathrm{m}^3/\mathrm{mol}$. Diamagnetic susceptibilities of

[14] J. Larmor, 1897.

the same order are found for a great number of elements as well. Experimental data obtained at room temperature for some of them are listed in Table 3.5.

Table 3.5. Molar and specific magnetic susceptibility values (in SI units) for diamagnetic elements at room temperatures. For carbon, diamond data are given

	χ_{mol} [SI] $\left(10^{-12}\,\text{m}^3/\text{mol}\right)$	χ_{mass} [SI] $\left(10^{-9}\,\text{m}^3/\text{kg}\right)$
Cu	−68.6	−1.08
Ag	−245	−2.27
Ga	−271	−3.9
C	−74.1	−6.17
Si	−39.2	−1.8
Ge	−146	−1.33
Sn	−470	−3.3
Sb	−1244	−10
Bi	−3520	−16.8
Se	−314	−4.0
Te	−478	−3.9

Alert readers have no doubt noticed that among diamagnetic materials there are not only covalently bonded semiconductors and insulators with closed electron shells but also semimetals, like bismuth, and a few metals, such as noble metals. In the latter electrons outside the closed d-shell can be considered practically free: they are responsible for electric conduction. While they give a positive, paramagnetic contribution to susceptibility,[15] the total susceptibility receives a more important diamagnetic contribution from d-electrons.

3.2.6 Atomic Paramagnetism

When outer shells are not completely filled and thus $J \neq 0$, the atom has a permanent magnetic moment. If an external magnetic field is applied, this magnetic moment will be more likely to line up with the field than against it. Consequently a magnetization proportional to the magnetic field appears. This is called *paramagnetic behavior*. The value of paramagnetic susceptibility can be easily determined classically. In a magnetic field of induction $\boldsymbol{B}$ the energy of a classical permanent magnetic moment $\boldsymbol{\mu}$ is $-\boldsymbol{\mu}\cdot\boldsymbol{B}$. Using classical Boltzmann statistics, the probability that the magnetic moment points in the elementary solid angle $\mathrm{d}\Omega$ around the direction given by the polar angles θ and φ is

[15] The susceptibility due to the electrons not bound to the ion core (free electrons) will be determined in Volume 2.

$$P(\theta,\varphi)\,\mathrm{d}\Omega = \frac{1}{Z}\mathrm{e}^{\boldsymbol{\mu}\cdot\boldsymbol{B}/k_\mathrm{B}T}\,\mathrm{d}\Omega\,, \tag{3.2.47}$$

where $\boldsymbol{\mu}\cdot\boldsymbol{B} = \mu B\cos\theta$, and Z is some normalization factor. With this weight factor, the thermal average of the magnetic moment is found to be

$$\langle\boldsymbol{\mu}\rangle = \frac{1}{Z}\int \boldsymbol{\mu}\,\mathrm{e}^{\boldsymbol{\mu}\cdot\boldsymbol{B}/k_\mathrm{B}T}\,\mathrm{d}\Omega\,. \tag{3.2.48}$$

If there are N magnetic moments in volume V, then magnetization is given by the density of the total magnetic moment, $\boldsymbol{M} = \langle\boldsymbol{\mu}\rangle N/V$, and susceptibility by its derivative with respect to magnetic field $\boldsymbol{H}$ – for which the approximation $\boldsymbol{H} = \boldsymbol{B}/\mu_0$ can be used. The components of the susceptibility tensor are thus

$$\begin{aligned}\chi_{\alpha\beta} &= \frac{N}{V}\frac{\mu_0}{k_\mathrm{B}T}\left\{\frac{1}{Z}\int \mu_\alpha\mu_\beta \mathrm{e}^{\boldsymbol{\mu}\cdot\boldsymbol{B}/k_\mathrm{B}T}\,\mathrm{d}\Omega\right.\\ &\qquad\left. -\frac{1}{Z^2}\int \mu_\alpha \mathrm{e}^{\boldsymbol{\mu}\cdot\boldsymbol{B}/k_\mathrm{B}T}\,\mathrm{d}\Omega\int \mu_\beta \mathrm{e}^{\boldsymbol{\mu}\cdot\boldsymbol{B}/k_\mathrm{B}T}\,\mathrm{d}\Omega\right\}\\ &= \frac{N}{V}\frac{\mu_0}{k_\mathrm{B}T}\big\{\langle\mu_\alpha\mu_\beta\rangle - \langle\mu_\alpha\rangle\langle\mu_\beta\rangle\big\}\,.\end{aligned} \tag{3.2.49}$$

Thermal fluctuations of the magnetic moment are recognized on the right-hand side. This is an example of the fluctuation–dissipation theorem which states that the linear response of the system to an external perturbation can be expressed in terms of the fluctuations of some quantity in thermal equilibrium.

The initial susceptibility is obtained by determining the thermal averages in (3.2.49) in thermal equilibrium, in the absence of any applied field. As free moments can point in any direction, $\langle\mu_\alpha\rangle = 0$, and the various components are uncorrelated – that is, off-diagonal elements vanish in $\langle\mu_\alpha\mu_\beta\rangle$. Equivalence of the three spatial directions implies that the diagonal elements are equal to one-third of the squared magnitude of the momentum,

$$\langle\mu_\alpha\mu_\beta\rangle = \tfrac{1}{3}\mu^2\delta_{\alpha\beta}\,. \tag{3.2.50}$$

Thus the magnetic susceptibility of a system made up of atoms with classical magnetic moments of magnitude μ is

$$\boxed{\chi = \frac{N}{V}\frac{\mu_0\mu^2}{3k_\mathrm{B}T}\,.} \tag{3.2.51}$$

It was first observed by P. Curie[16] (1895) that the susceptibility of paramagnetic materials varies inversely with temperature. This temperature dependence is called *Curie's law*, and susceptibilities proportional to $1/T$ are

[16] Pierre Curie (1859–1906) shared the Nobel Prize with his wife, Marie Curie née Sklodowska (1867–1934), and Henri Becquerel (1852–1908); the Curies received the accolade "in recognition of the extraordinary services they have rendered by their joint researches on the radiation phenomena discovered by Professor Henri Becquerel".

in general known as *Curie susceptibilities*. The above formula was, however, first derived by P. Langevin (1905), hence the name *Langevin susceptibility* is also used.

To obtain a more precise, quantum mechanical description of paramagnetic behavior, one has to focus on the second term in the Hamiltonian (3.2.27),

$$\mu_{\mathrm{B}}\,(\boldsymbol{L} + |g_{\mathrm{e}}|\boldsymbol{S}) \cdot \boldsymbol{B}\,. \tag{3.2.52}$$

Assuming that the applied magnetic field is weak, the contribution of this term is calculated in the first nonvanishing order of perturbation theory. For this one needs to know the matrix elements of $\boldsymbol{L}$ and $\boldsymbol{S}$ in the atomic state.

Because of spin–orbit interaction, $\boldsymbol{L}$ and $\boldsymbol{S}$ are not conserved separately: the state is determined by the magnitude of the resultant $\boldsymbol{J} = \boldsymbol{L} + \boldsymbol{S}$ and its component along the magnetic field. When the z-axis is chosen along the magnetic field direction, states are characterized by the quantum numbers L, S, J, and m_J – the last being the eigenvalue of J_z. The matrix elements necessary for determining the energy are obtained using the *Wigner–Eckart theorem*.[17] Before formulating this theorem, the concept of *vector operators* has to be introduced first.

An operator $\boldsymbol{V}$ is called a vector operator if it is transformed as a vector under rotations – or, equivalently, if its commutation relations with the components of the total angular momentum are

$$\begin{aligned} [J_x, V_x] &= 0\,, \\ [J_x, V_y] &= \mathrm{i}\hbar V_z\,, \\ [J_x, V_z] &= -\mathrm{i}\hbar V_y\,, \end{aligned} \tag{3.2.53}$$

and other similar expressions obtained by cyclic permutation of the indices. Making use of the equality $\boldsymbol{J} = \boldsymbol{L} + \boldsymbol{S}$, it is easy to show that both $\boldsymbol{L}$ and $\boldsymbol{S}$ satisfy the above conditions, i.e., they are vector operators.

The Wigner–Eckart theorem asserts that for fixed L and S, in the $(2L+1)(2S+1)$-dimensional subspace of the eigenfunctions of the total angular momentum operator, the matrix elements of any vector operator are proportional to the matrix elements of the total angular momentum operator, and the constant of proportionality is independent of the value of m_J:

$$\langle L, S, J, m_J|\boldsymbol{V}|L, S, J, m'_J\rangle = \alpha(L, S, J)\langle L, S, J, m_J|\boldsymbol{J}|L, S, J, m'_J\rangle\,. \tag{3.2.54}$$

In other words: within this subspace the vector operator $\boldsymbol{V}$ is equivalent to the quantity $\alpha\boldsymbol{J}$:

$$\boldsymbol{V} \equiv \alpha(L, S, J)\boldsymbol{J}\,. \tag{3.2.55}$$

[17] C. Eckart (1930), E. P. Wigner (1931). Eugene Paul Wigner (1902–1995) was awarded the Nobel Prize in 1963 "for his contributions to the theory of the atomic nucleus and the elementary particles, particularly through the discovery and application of fundamental symmetry principles".

Let us now introduce the operator

$$P(L,S,J) = \sum_{m_J} |L,S,J,m_J\rangle\langle L,S,J,m_J| \,. \tag{3.2.56}$$

Exploiting the mutual orthogonality of angular momentum eigenfunctions it is readily seen that $P^2(L,S,J) = P(L,S,J)$, and also that the completeness relation

$$\sum_J P(L,S,J) = \sum_J \sum_{m_J} |L,S,J,m_J\rangle\langle L,S,J,m_J| = 1 \tag{3.2.57}$$

holds, thus for given L and S the operator $P(L,S,J)$ projects onto the subspace of quantum number J. Moreover, this projection operator commutes will all three components of $\boldsymbol{J}$. When (3.2.54) is multiplied by $|L,S,J,m_J\rangle$ from the left and by $\langle L,S,J,m'_J|$ from the right, summation yields

$$P(L,S,J)\boldsymbol{V}P(L,S,J) = \alpha(L,S,J)P(L,S,J)\boldsymbol{J}P(L,S,J) \,. \tag{3.2.58}$$

To determine the coefficient $\alpha(L,S,J)$, consider the following matrix element of the scalar $\boldsymbol{V}\cdot\boldsymbol{J}$:

$$\langle L,S,J,m_J|\boldsymbol{V}\cdot\boldsymbol{J}|L,S,J',m_{J'}\rangle. \tag{3.2.59}$$

Making use of the completeness relation, the fact that only those matrix elements of $\boldsymbol{J}$ are nonzero that belong to the same quantum number J, and the Wigner–Eckart theorem we have

$$\langle L,S,J,m_J|\boldsymbol{V}\cdot\boldsymbol{J}|L,S,J',m_{J'}\rangle = \alpha(L,S,J)\langle L,S,J,m_J|\boldsymbol{J}^2|L,S,J',m_{J'}\rangle. \tag{3.2.60}$$

When $\alpha(L,S,J)$ is expressed, we arrive at

$$\boldsymbol{V} = \frac{\langle \boldsymbol{V}\cdot\boldsymbol{J}\rangle}{\langle \boldsymbol{J}^2\rangle}\,\boldsymbol{J} \,. \tag{3.2.61}$$

in the space of the angular momentum eigenfunctions. Thus one can draw the intuitive conclusion that within this subspace only the $\boldsymbol{J}$-projection of the vector operator is important.

For the vector operator $\boldsymbol{L} + |g_\mathrm{e}|\boldsymbol{S}$ this implies

$$\boldsymbol{L} + |g_\mathrm{e}|\boldsymbol{S} = g_J\,\boldsymbol{J} \tag{3.2.62}$$

in the relevant subspace. The value of g_J, called the *Landé g-factor*[18] can be determined from

$$g_J = \frac{\langle(\boldsymbol{L} + |g_\mathrm{e}|\boldsymbol{S})\cdot\boldsymbol{J}\rangle}{\langle \boldsymbol{J}^2\rangle}. \tag{3.2.63}$$

[18] A. Landé, 1923.

One only has to make use of

$$\boldsymbol{L}\cdot\boldsymbol{S}=\tfrac{1}{2}[\boldsymbol{J}^2-\boldsymbol{L}^2-\boldsymbol{S}^2]\,, \tag{3.2.64}$$

and thus

$$\boldsymbol{L}\cdot\boldsymbol{J}=\boldsymbol{L}\cdot(\boldsymbol{L}+\boldsymbol{S})=\boldsymbol{L}^2+\tfrac{1}{2}[\boldsymbol{J}^2-\boldsymbol{L}^2-\boldsymbol{S}^2]\,, \tag{3.2.65}$$

as well as

$$\boldsymbol{S}\cdot\boldsymbol{J}=\boldsymbol{S}\cdot(\boldsymbol{L}+\boldsymbol{S})=\boldsymbol{S}^2+\tfrac{1}{2}[\boldsymbol{J}^2-\boldsymbol{L}^2-\boldsymbol{S}^2]\,. \tag{3.2.66}$$

In the subspace of the eigenvectors of $\boldsymbol{J}=\boldsymbol{L}+\boldsymbol{S}$

$$\begin{aligned}\langle\boldsymbol{L}\cdot\boldsymbol{J}\rangle&=L(L+1)+\tfrac{1}{2}[J(J+1)-L(L+1)-S(S+1)]\,,\\ \langle\boldsymbol{S}\cdot\boldsymbol{J}\rangle&=S(S+1)+\tfrac{1}{2}[J(J+1)-L(L+1)-S(S+1)]\,.\end{aligned} \tag{3.2.67}$$

This finally yields

$$\begin{aligned}g_J&=\tfrac{1}{2}(|g_\mathrm{e}|+1)+\tfrac{1}{2}(|g_\mathrm{e}|-1)\frac{S(S+1)-L(L+1)}{J(J+1)}\\ &=1+(|g_\mathrm{e}|-1)\frac{J(J+1)+S(S+1)-L(L+1)}{2J(J+1)}\,.\end{aligned} \tag{3.2.68}$$

The importance of this result lies in the fact that g_J determines the total magnetic moment

$$\boldsymbol{\mu}=-\mu_\mathrm{B}(\boldsymbol{L}+|g_\mathrm{e}|\boldsymbol{S})=-g_J\mu_\mathrm{B}\boldsymbol{J} \tag{3.2.69}$$

of an ion core, and therefore the energy level splitting in an applied magnetic field.

The interaction of the magnetic moment with an applied magnetic field is included in the Hamiltonian as the perturbation term

$$\mathcal{H}_\mathrm{ext}=-\boldsymbol{\mu}\cdot\boldsymbol{B}=g_J\mu_\mathrm{B}\,\boldsymbol{J}\cdot\boldsymbol{B}\,. \tag{3.2.70}$$

This completely lifts the $2J+1$-fold degeneracy of the original state, splitting it into $2J+1$ nondegenerate Zeeman levels. When the quantization axis is chosen along the applied field direction, the energy shift of the individual levels is given by

$$\Delta E=g_J\mu_\mathrm{B}m_J B\,, \tag{3.2.71}$$

where $m_J=-J,-J+1,\ldots,J$.

A simple physical interpretation can be given to the foregoing by picturing the moments $\boldsymbol{L}$ and $\boldsymbol{S}$ as classical vectors of length $\sqrt{L(L+1)}$ and $\sqrt{S(S+1)}$. Because of the spin–orbit interaction, these will precess about a third classical vector, $\boldsymbol{J}$, of length $\sqrt{J(J+1)}$, as shown in Fig. 3.1.

Owing to the precession, all components perpendicular to the direction of $\boldsymbol{J}$ are averaged out, and only the projection of $\boldsymbol{L}$ and $\boldsymbol{S}$ on the direction of $\boldsymbol{J}$ survive. As the gyromagnetic ratio is different for spin and orbital angular

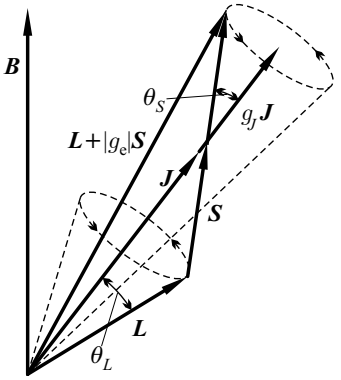


Fig. 3.1. Precession of the orbital angular momentum and the spin about the total angular momentum

momentum, the total magnetic moment – that is, the resultant of the intrinsic magnetic moment and the magnetic moment associated with the orbital angular momentum – does not point in the same direction as $\boldsymbol{J}$ but rather precesses about it, and only its component parallel to $\boldsymbol{J}$ is conserved. This component is given in terms of the lengths of the vectors $\boldsymbol{\mu}_L$ and $\boldsymbol{\mu}_S$,

$$|\boldsymbol{\mu}_L| = \mu_\mathrm{B}\sqrt{L(L+1)}\,, \qquad |\boldsymbol{\mu}_S| = |g_\mathrm{e}|\mu_\mathrm{B}\sqrt{S(S+1)} \tag{3.2.72}$$

and the angles θ_L and θ_S that $\boldsymbol{J}$ makes with $\boldsymbol{L}$ and $\boldsymbol{S}$ as

$$|\boldsymbol{\mu}_J| = |\boldsymbol{\mu}_L|\cos\theta_L + |\boldsymbol{\mu}_S|\cos\theta_S\,. \tag{3.2.73}$$

By reading off the angles from the figure and making use of the relation

$$|\boldsymbol{\mu}_J| = g_J\sqrt{J(J+1)}\mu_\mathrm{B} \tag{3.2.74}$$

the previous result for the Landé g-factor is recovered. When a magnetic field is turned on, this average magnetic moment (of magnitude $|\boldsymbol{\mu}_J|$ and direction $\boldsymbol{J}$) starts to precess about the field direction, and only its component along the field contributes to energy. The classical picture cannot account for the fact that the component along the field direction is quantized, i.e., it can take only discrete values.

To specify the Curie or Langevin susceptibility, we shall assume that the energy difference between the ground state (as determined by Hund's rules) and higher-lying states is larger than the thermal energy $k_\mathrm{B}T$ at experimentally relevant temperatures, and consequently the thermal occupation of the latter states can be neglected. If this assumption is justified, only a single level with quantum number J has to be considered for each atom – which, however, splits into $2J+1$ levels in a magnetic field. A well-known theorem in statistical mechanics states that the partition function of a system made

up of independent subsystems is the product of the partition functions of the subsystems. In a first approximation, the state of individual atoms will be considered independent of each another. The partition function Z_a of all possible states of a single atom is given by

$$Z_\mathrm{a} = \sum_n \mathrm{e}^{-\beta E_n} = \sum_{m_J=-J}^{J} \exp(-\beta E_0) \exp(-\beta g_J \mu_\mathrm{B} m_J B)\,. \tag{3.2.75}$$

Summation yields

$$\begin{aligned} Z_\mathrm{a} &= \exp(-\beta E_0) \exp(-\beta g_J \mu_\mathrm{B} J B)\, \frac{\exp[\beta g_J \mu_\mathrm{B}(2J+1)B] - 1}{\exp(\beta g_J \mu_\mathrm{B} B) - 1} \\ &= \mathrm{e}^{-\beta E_0}\, \frac{\exp[\beta g_J \mu_\mathrm{B}(J+1/2)B] - \exp[-\beta g_J \mu_\mathrm{B}(J+1/2)B]}{\exp(\beta g_J \mu_\mathrm{B} B/2) - \exp(-\beta g_J \mu_\mathrm{B} B/2)} \\ &= \mathrm{e}^{-\beta E_0}\, \frac{\sinh[\beta g_J \mu_\mathrm{B}(J+1/2)B]}{\sinh[\beta g_J \mu_\mathrm{B} B/2]}\,. \end{aligned} \tag{3.2.76}$$

Because of the independence of individual atoms, the partition function of the entire system with N atoms is just the Nth power of the partition function of a single atom,

$$Z = Z_\mathrm{a}^N\,. \tag{3.2.77}$$

As free energy is the logarithm of the partition function, the component of magnetization along the magnetic field is

$$M = -\frac{1}{V}\frac{\partial F}{\partial B} = \frac{N}{V} g_J \mu_\mathrm{B}\, J B_J(\beta g_J \mu_\mathrm{B} J B)\,, \tag{3.2.78}$$

where $B_J(x)$, the *Brillouin function* can be written as

$$B_J(x) = \frac{2J+1}{2J} \coth \frac{2J+1}{2J} x - \frac{1}{2J} \coth \frac{1}{2J} x\,. \tag{3.2.79}$$

For $J = 1/2$, the Brillouin function takes the particularly simple form

$$B_{1/2}(x) = \tanh x\,. \tag{3.2.80}$$

In the $J \gg 1$ limit the component of the magnetic moment along the magnetic field is found to be

$$\langle \mu_z \rangle = \mu\, L(\mu B / k_\mathrm{B} T)\,, \tag{3.2.81}$$

in perfect agreement with the result that would be obtained from the classical equation (3.2.48). Here

$$L(x) = \coth x - 1/x \tag{3.2.82}$$

is the *Langevin function*.

In strong fields, where the energy due to the coupling to the magnetic field is much larger than the thermal energy ($g_J \mu_B JB \gg k_B T$), the projection of the spins will be maximal along the field direction, and so magnetization reaches saturation. Apart from extremely low temperatures, $\mu_B B \ll k_B T$ in fields customarily applied in susceptibility measurements, therefore the argument of the Brillouin function is small. Using the small-x expansion of the coth function,

$$\coth x = \frac{1}{x} + \frac{1}{3}x - \frac{1}{45}x^3 + \ldots , \tag{3.2.83}$$

we have

$$B_J(x) = \frac{J+1}{J}\,\frac{x}{3} - \frac{(2J+1)^4 - 1}{(2J)^4}\,\frac{x^3}{45} + \ldots . \tag{3.2.84}$$

Substitution into (3.2.78) gives the leading-order term of the susceptibility:

$$\boxed{\chi = \frac{N}{V}\,\frac{\mu_0 (g_J \mu_B)^2 J(J+1)}{3 k_B T}\,.} \tag{3.2.85}$$

To compare this formula with the expression in (3.2.51) valid for classical magnetic moments, the *effective magnetic moment* μ_{eff} and the *effective magneton number* p are customarily introduced via

$$\chi = \frac{N}{V}\,\frac{\mu_0 \mu_{\text{eff}}^2}{3 k_B T} = \frac{N}{V}\,\frac{\mu_0 \mu_B^2 p^2}{3 k_B T}\,. \tag{3.2.86}$$

Their theoretical values are $\mu_{\text{eff}}^{\text{th}} = g_J \sqrt{J(J+1)}\,\mu_B$ and $p^{\text{th}} = g_J \sqrt{J(J+1)}$. Table 3.6 lists the values of p calculated from Hund's rules for trivalent rare-earth metals (p^{th}) as well as determined from experimental data on the susceptibility of paramagnetic salts containing such ions (p^{exp}).

As the data in the table show, the agreement between calculated and measured values is excellent – except for samarium and europium. The reason for the discrepancy in the latter is that our earlier assumption – that it is sufficient to consider only the lowest $2J + 1$ energy levels associated with the value of J given by Hund's rules – is not justified now. In this case somewhat higher-lying states must also be included in calculations of the susceptibility, as their occupation can no longer be neglected. We shall see a remarkable example in the following subsection.

A far worse agreement is found for 3d transition metals when the calculated magnetic moment of the configuration given by Hund's rules is compared to the effective moment determined from measured data of the paramagnetic susceptibility in solids containing transition-metal ions. This is illustrated in Table 3.7.

The table clearly shows that an agreement can be found only when the effective magneton number due to the electrons' intrinsic magnetic moment

Table 3.6. Ground-state electron configuration, Landé g-factor, as well as calculated and measured effective magneton numbers for trivalent rare-earth ions

Ion	$4f^n$	L	L	J	Ground state	g_J	p^{th}	p^{exp}
La^{3+}	0	0	0	0	$^1S_0$	0	0	0
Ce^{3+}	1	1/2	3	5/2	$^2F_{5/2}$	6/7	2.535	2.4–2.7
Pr^{3+}	2	1	5	4	$^3H_4$	4/5	3.578	3.4–3.6
Nd^{3+}	3	3/2	6	9/2	$^4I_{9/2}$	8/11	3.618	3.4–3.7
Pm^{3+}	4	2	6	4	$^5I_4$	3/5	2.683	–
Sm^{3+}	5	5/2	5	5/2	$^6H_{5/2}$	2/7	0.845	1.3–1.6
Eu^{3+}	6	3	3	0	$^7F_0$	0	0	3.2–3.4
Gd^{3+}	7	7/2	0	7/2	$^8S_{7/2}$	2	7.937	7.9–8.0
Tb^{3+}	8	3	3	6	$^7F_6$	3/2	9.721	9.4–9.8
Dy^{3+}	9	5/2	5	15/2	$^6H_{15/2}$	4/3	10.646	10.5–10.7
Ho^{3+}	10	2	6	8	$^5I_8$	5/4	10.607	10.3–10.6
Er^{3+}	11	3/2	6	15/2	$^4I_{15/2}$	6/5	9.581	9.4–9.6
Tm^{3+}	12	1	5	6	$^3H_6$	7/6	7.561	7.3–7.6
Yb^{3+}	13	1/2	3	7/2	$^2F_{7/2}$	8/7	4.536	4.4–4.6
Lu^{3+}	14	0	0	0	$^1S_0$	0	0	0

Table 3.7. Electron configuration, quantum numbers of the ground state (as specified by Hund's rules), calculated and measured values of the effective magneton number p, and spin the theoretically determined contribution to the effective magneton number p_{s} for some transition-metal ions

Ion	$3d^n$	S	L	J	Ground state	g_J	p^{th}	p^{exp}	$p_{\mathrm{s}}^{\mathrm{th}}$
Ti^{3+}	1	1/2	2	3/2	$^2D_{3/2}$	4/5	1.549	1.6–1.8	1.732
V^{4+}	1	1/2	2	3/2	$^2D_{3/2}$	4/5	1.549	1.6–1.8	1.732
V^{3+}	2	1	3	2	$^3F_2$	2/3	1.633	2.5–2.8	2.828
V^{2+}	3	3/2	3	3/2	$^4F_{3/2}$	2/5	0.775	3.7–4.0	3.873
Cr^{3+}	3	3/2	3	3/2	$^4F_{3/2}$	2/5	0.775	3.7–4.0	3.873
Mn^{4+}	3	3/2	3	3/2	$^4F_{3/2}$	2/5	0.775	3.7–4.0	3.873
Cr^{2+}	4	2	2	0	$^5D_0$	–	0	4.8–5.0	4.900
Mn^{3+}	4	2	2	0	$^5D_0$	–	0	4.8–5.0	4.900
Mn^{2+}	5	5/2	0	5/2	$^6S_{5/2}$	2	5.916	5.9	5.916
Fe^{3+}	5	5/2	0	5/2	$^6S_{5/2}$	2	5.916	5.9	5.916
Fe^{2+}	6	2	2	4	$^5D_4$	3/2	6.708	5.2–5.4	4.900
Co^{3+}	6	2	2	4	$^5D_4$	3/2	6.708	5.2–5.4	4.900
Co^{2+}	7	3/2	3	9/2	$^4F_{9/2}$	4/3	6.633	4.8–4.9	3.873
Ni^{2+}	8	1	3	4	$^3F_4$	5/4	5.590	3.1–3.2	2.828
Cu^{2+}	9	1/2	2	5/2	$^2D_{5/2}$	6/5	3.550	1.9	1.732
Zn^{2+}	10	0	0	0	$^1S_0$	0	0	0	0

$p_\mathrm{s} = |g_\mathrm{e}|\sqrt{S(S+1)}$ is considered, i.e., the contribution of the orbital angular momentum is neglected. It seems as if the electrons' orbit were rigidly fixed – quenched – and unaffected by the magnetic field. As we shall see in Chapter 6, this happens because inside the solid the transition-metal ions feel the relatively strong electrostatic field of the neighboring ions. Therefore ions can no longer be considered to be independent of each other, and so Hund's rules may become invalid. For iron, cobalt, and nickel ions the agreement is not very good even for p_s when the effective moment is taken from room-temperature measurements. The reason for this is that for such ions thermal energies are on the order of the energy difference between the ground state and higher-lying energy levels, and thus the contribution of the latter to susceptibility cannot be neglected.

3.2.7 Van Vleck Paramagnetism

In this subsection we shall return to the problem of europium, in which the outermost electron shell is not closed although according to Hund's rules $J = 0$ in the ground state. The same situation occurs when two electrons are present in the p-shell or four in the d-shell (or, in general, $n = 2l$ electrons in the shell with azimuthal quantum number l). In such cases the first nonvanishing contribution to paramagnetic susceptibility is obtained in the second order of perturbation theory, and so higher-lying energy levels must also be taken into account as intermediate states. The second-order correction to the ground-state energy is given by

$$\Delta E_0 = \sum_{n\neq 0} \frac{|\langle 0|\mu_\mathrm{B} B(L_z + |g_\mathrm{e}|S_z)|n\rangle|^2}{E_0 - E_n}\,. \tag{3.2.87}$$

This energy shift is always negative as the energy denominator is negative in each term. The resulting magnetization is always along the magnetic field direction, and the susceptibility is positive:

$$\chi = 2\frac{N}{V}\mu_0\mu_\mathrm{B}^2 \sum_n \frac{|\langle 0|(L_z + |g_\mathrm{e}|S_z)|n\rangle|^2}{E_n - E_0}\,. \tag{3.2.88}$$

This phenomenon is called *Van Vleck paramagnetism*, and χ the *Van Vleck susceptibility*.[19] If the excitation energy of the state $|n\rangle$ is not too large, the small energy denominator may make this correction more important than the diamagnetic contribution. When this Van Vleck contribution is taken into account, the calculated paramagnetic susceptibility of substances containing trivalent europium ions is in good agreement with measured data.

In the ground state of the samarium ion J takes a nonzero value, therefore Sm^{3+} has a nonvanishing paramagnetic moment. Nevertheless the lowest-lying excited states are close enough for that the Van Vleck second-order correction

[19] J. H. Van Vleck, 1932. See footnote on page 6.

be comparable to the paramagnetic contribution. This explains why the effective magneton number determined from the ground-state configuration does not agree with the measured value.

3.2.8 Electron Spin Resonance

As we have seen, when a paramagnetic ion of angular momentum $\boldsymbol{J}$ is placed in a uniform magnetic field, the initial $2J+1$-fold degeneracy of the ground state – that is due to the $2J+1$ possible values of the quantum number m_J – is split. In the presence of an electromagnetic radiation, transitions between neighboring levels become possible in accordance with the selection rule $\Delta m_J = \pm 1$. As the energy difference between two adjacent Zeeman levels is

$$\Delta E = g_J \mu_\mathrm{B} B \,, \tag{3.2.89}$$

atoms can absorb photons of this energy, and thus the electromagnetic field of the corresponding frequency is absorbed in a resonant manner. This phenomenon is called *electron spin resonance* (ESR) or *electron paramagnetic resonance* (EPR). In fields on the order of several teslas, this resonant absorption occurs at a frequency of a few GHz (or, in terms of the wavelength, in the centimeter (microwave) range). In experimental studies samples are placed in a microwave field whose frequency is determined by the dimensions of the cavity, and absorption is measured as the strength of the applied field is varied. Then the Landé g_J-factor is determined from the transition frequency. The method is therefore suited to identifying paramagnetic ions within the sample.

To understand resonant absorption, consider the magnetic moment $\boldsymbol{\mu}_J$ as a classical vector that has an angular momentum $\hbar \boldsymbol{J} = -\boldsymbol{\mu}_J/\gamma_J$, where

$$\gamma_J = g_J \frac{\mu_\mathrm{B}}{\hbar} = g_J \frac{e}{2m_\mathrm{e}} \tag{3.2.90}$$

is the gyromagnetic ratio. For notational simplicity, we shall omit the label J. A magnetic moment $\boldsymbol{\mu}$ whose axis makes an angle θ with the direction of the uniform magnetic field $\boldsymbol{B}$ experiences a torque

$$\boldsymbol{M} = \boldsymbol{\mu} \times \boldsymbol{B} \,, \tag{3.2.91}$$

therefore the equation of motion of the magnetic moment is given by

$$\boxed{\frac{1}{\gamma}\frac{\mathrm{d}\boldsymbol{\mu}}{\mathrm{d}t} = \boldsymbol{B} \times \boldsymbol{\mu} \,.} \tag{3.2.92}$$

Thus the moment precesses about the magnetic field direction at a constant angular frequency $\omega_\mathrm{L} = \gamma B$ – the *Larmor frequency*. When a weak perpendicular field of angular frequency ω is superimposed on the uniform field, the precessing moment will resonantly absorb energy from the magnetic

field, provided ω is equal to the precession frequency ω_{L}. For all other frequencies the relative phases shift continuously with time, and so no energy will be absorbed.

For a quantum mechanical derivation of this equation we start with

$$\frac{\mathrm{d}\boldsymbol{\mu}(t)}{\mathrm{d}t} = \frac{\mathrm{i}}{\hbar}[\mathcal{H}(t), \boldsymbol{\mu}(t)]\,, \tag{3.2.93}$$

where the contribution of the magnetic field to the total Hamiltonian is

$$\mathcal{H}_{\mathrm{ext}} = -\boldsymbol{\mu} \cdot \boldsymbol{B}(t)\,. \tag{3.2.94}$$

According to (3.2.69), the magnetic moment is proportional to the total angular momentum, thus the known commutation relations of the latter lead to

$$\frac{\mathrm{d}\boldsymbol{\mu}}{\mathrm{d}t} = \gamma \boldsymbol{B} \times \boldsymbol{\mu}\,. \tag{3.2.95}$$

When quantum mechanical averages of the operators are taken, classical equations of motions are recovered. In what follows, this average will be meant by $\boldsymbol{\mu}$.

We shall now consider the previously outlined case. Two fields are applied: a constant field, of magnitude B_0, along the z-axis, and a linearly polarized time-dependent one, of amplitude $2B_1$ and angular frequency ω, along the x-axis:

$$B_x = 2B_1 \cos\omega t\,, \qquad B_y = 0\,, \qquad B_z = B_0\,. \tag{3.2.96}$$

The equation of motion in component form is

$$\begin{aligned} \frac{\mathrm{d}\mu_x}{\mathrm{d}t} &= -\gamma \mu_y B_0\,, \\ \frac{\mathrm{d}\mu_y}{\mathrm{d}t} &= \gamma \mu_x B_0 - 2\gamma \mu_z B_1 \cos\omega t\,, \\ \frac{\mathrm{d}\mu_z}{\mathrm{d}t} &= 2\gamma \mu_y B_1 \cos\omega t\,. \end{aligned} \tag{3.2.97}$$

If the time-dependent field is assumed to be weak compared to the constant field, then the time dependence of the z component of the magnetic moment can be ignored, as it is on the order of B_1^2. The solution of the system of equations is then

$$\mu_x = \frac{2\gamma B_1\, \omega_{\mathrm{L}} \langle \mu_z \rangle}{\omega_{\mathrm{L}}^2 - \omega^2} \cos\omega t\,, \qquad \mu_y = \frac{2\gamma B_1\, \omega \langle \mu_z \rangle}{\omega_{\mathrm{L}}^2 - \omega^2} \sin\omega t\,, \tag{3.2.98}$$

where $\omega_{\mathrm{L}} = \gamma B_0$. Besides μ_x (which is in phase with the perturbing field) a y component (which lags $\pi/2$ behind the applied field) is also present. As a generalization of this result the time dependence of the magnetic moment can be written as

$$\boldsymbol{\mu} = \langle \boldsymbol{\mu} \rangle + \boldsymbol{\chi}' \frac{2\boldsymbol{B}_1}{\mu_0} \cos \omega t + \boldsymbol{\chi}'' \frac{2\boldsymbol{B}_1}{\mu_0} \sin \omega t \,, \tag{3.2.99}$$

where the quantities $\boldsymbol{\chi}'$ and $\boldsymbol{\chi}''$ are tensors. The term proportional to $\boldsymbol{\chi}'$ gives the response that is in phase with the applied field, so $\boldsymbol{\chi}'$ is the usual susceptibility. The term proportional to $\boldsymbol{\chi}''$, which is $\pi/2$ out of phase with the applied field, can be related to energy absorption, as the energy absorbed per unit time is

$$Q = \frac{\omega}{2\mu_0} \boldsymbol{B}_1 \, \boldsymbol{\chi}'' \, \boldsymbol{B}_1 \,. \tag{3.2.100}$$

It is customary to consider $\boldsymbol{\chi}'$ and $\boldsymbol{\chi}''$ as the real and imaginary parts of a complex susceptibility, $\boldsymbol{\chi} = \boldsymbol{\chi}' + \mathrm{i}\boldsymbol{\chi}''$. Then

$$\boldsymbol{\mu} = \langle \boldsymbol{\mu} \rangle + \boldsymbol{\chi} \frac{\boldsymbol{B}_1}{\mu_0} \mathrm{e}^{-\mathrm{i}\omega t} + \boldsymbol{\chi}^* \frac{\boldsymbol{B}_1}{\mu_0} \mathrm{e}^{\mathrm{i}\omega t} \,. \tag{3.2.101}$$

The linearly polarized field can be made up of two counterrotating circularly polarized radiations; in this case $\boldsymbol{\chi}$ and its complex conjugate correspond to the response to positive and negative polarization, respectively. From the expressions in (3.2.98) it is readily seen that

$$\chi'_{xx} = \chi''_{yx} = \frac{\gamma \mu_0 \omega_\mathrm{L} \langle \mu_z \rangle}{\omega_\mathrm{L}^2 - \omega^2} = \frac{\omega_\mathrm{L}^2}{\omega_\mathrm{L}^2 - \omega^2} \chi_0 \,, \tag{3.2.102}$$

where $\chi_0 = \mu_0 \langle \mu_z \rangle / B_0$ is the paramagnetic susceptibility.

The above expressions are singular at $\omega = \omega_\mathrm{L}$. The imaginary part of χ_{xx} cannot be precisely determined from these equations. One can only say that it vanishes for all frequencies $\omega \neq \omega_\mathrm{L}$, i.e., energy absorption is possible only at the resonance frequency. The amplitude can nevertheless be determined when one takes into account the interrelatedness of the real and complex parts. To see this, consider the so-called relaxation function, $\boldsymbol{\phi}(t - t_0)$. This describes how a system in thermal equilibrium in a magnetic field of unit strength will relax toward its new equilibrium state when the magnetic field is suddenly turned off at time t_0. Similarly, $-\boldsymbol{\phi}'(t - t_0) = -\mathrm{d}\boldsymbol{\phi}(t - t_0)/\mathrm{d}t$ gives the magnetization caused at time t by the application of a delta function pulse at an earlier time t_0. For time-dependent magnetic fields integration of the contributions from times prior to t leads to

$$\boldsymbol{\mu}(t) = \langle \boldsymbol{\mu} \rangle - \int_{-\infty}^{t} \boldsymbol{\phi}'(t - t_0) \boldsymbol{B}(t_0) \, \mathrm{d}t_0 \,. \tag{3.2.103}$$

When a harmonic magnetic field $B(t) = \mu_0 H \mathrm{e}^{-\mathrm{i}\omega t}$ of frequency ω is applied, the time-dependent part of the magnetic moment will vary with the same frequency,

$$\boldsymbol{\mu}(t) = \langle \boldsymbol{\mu} \rangle + \boldsymbol{\mu}(\omega) \mathrm{e}^{-\mathrm{i}\omega t} \,, \tag{3.2.104}$$

and thus the frequency-dependent magnetic susceptibility is given by

$$\boldsymbol{\chi}(\omega) = -\mu_0 \int_{-\infty}^{t} \boldsymbol{\phi}'(t-t_0)\mathrm{e}^{\mathrm{i}\omega(t-t_0)}\mathrm{d}t_0 = -\mu_0 \int_{0}^{\infty} \boldsymbol{\phi}'(t)\mathrm{e}^{\mathrm{i}\omega t}\mathrm{d}t\,. \qquad (3.2.105)$$

Integration by parts now leads to

$$\boldsymbol{\chi}(\omega) = \boldsymbol{\chi}_0 + \mathrm{i}\omega\,\mu_0 \int_{0}^{\infty} \boldsymbol{\phi}(t)\mathrm{e}^{\mathrm{i}\omega t}\mathrm{d}t\,. \qquad (3.2.106)$$

One may, in principle, obtain the relaxation function – and thus the real and imaginary parts of the susceptibility as well – by solving the earlier equations for the temporal variation of the magnetic moment in the special case when the magnetic field is suddenly turned off. Instead of proceeding along this line we shall make use of only one point of the above considerations: that magnetization is the causal response to an applied field. As it will be demonstrated in the Appendix of Volume 3, causality implies that when the frequency variable is extended to the complex plane, the frequency-dependent susceptibility is analytic on the upper half plane – moreover, its real and imaginary parts are related by the *Kramers–Kronig relations*[20]

$$\chi'(\omega) = \chi'(\infty) + \frac{2}{\pi}\int_{0}^{\infty} \frac{\omega'\chi''(\omega')}{\omega'^2 - \omega^2}\,\mathrm{d}\omega'\,, \qquad (3.2.107\text{-a})$$

$$\chi''(\omega) = -\frac{2}{\pi}\int_{0}^{\infty} \frac{\omega\chi'(\omega')}{\omega'^2 - \omega^2}\,\mathrm{d}\omega'\,. \qquad (3.2.107\text{-b})$$

Using the previously derived form for χ'_{xx}, we have

$$\chi''_{xx} = \frac{\pi}{2}\delta(\omega - \omega_\mathrm{L})\omega_\mathrm{L}\,\chi_0\,. \qquad (3.2.108)$$

A sharp resonance appears only because we have ignored all mechanisms that can cause the magnetic moment to relax. For an isolated atom the z component of the angular momentum is conserved. However, in condensed matter angular momentum can be transferred to other atoms or the lattice, and this leads to the relaxation of the magnetic moment toward its equilibrium value. This phenomenon is called the *spin–lattice relaxation*. We shall not discuss these interactions in detail here – nonetheless assume that the relaxation of a magnetic moment can be characterized by a *longitudinal* or *spin–lattice relaxation time* T_1. Then the equation of motion for the z component of the magnetization is

$$\frac{\mathrm{d}\mu_z}{\mathrm{d}t} = 2\gamma\mu_y B_1 \cos\omega t - \frac{\mu_z - \langle\mu_z\rangle}{T_1}\,. \qquad (3.2.109)$$

[20] H. A. Kramers, 1927, R. Kronig, 1926.

The new term corresponds to the assumption that without the driving precessing field the system would approach its equilibrium state exponentially, with a characteristic time T_1. As the variation of the spin component along the field $\boldsymbol{B}_0$ is accompanied by the variation of the magnetic energy of the ion, during the spin–lattice relaxation process energy has to be transferred to another part of the system: the lattice.

Other ions also influence the temporal variations of the x and y components of the magnetic moment, as the moment is moving in the nonuniform field created by its neighbors. This can be characterized phenomenologically by a *transverse* or *spin–spin relaxation time* T_2. As no energy transfer occurs during transverse relaxation, spin–spin relaxation is usually several orders of magnitude faster than spin–lattice relaxation: $T_2 \ll T_1$. Typical values are $T_1 \sim 10^{-6}\,\mathrm{s}$ and $T_2 \sim 10^{-10}\,\mathrm{s}$.

When relaxation is taken into account, the equation of motion reads

$$\begin{aligned} \frac{\mathrm{d}\mu_x}{\mathrm{d}t} &= -\gamma\mu_y B_0 - \frac{\mu_x}{T_2}\,, \\ \frac{\mathrm{d}\mu_y}{\mathrm{d}t} &= \gamma\mu_x B_0 - 2\gamma\mu_z B_1 \cos\omega t - \frac{\mu_y}{T_2}\,. \end{aligned} \tag{3.2.110}$$

These equations and (3.2.109) are jointly called the *Bloch equations*.[21] When only the constant field is present, the previous equations simplify to

$$\begin{aligned} \frac{\mathrm{d}\mu_x}{\mathrm{d}t} &= -\gamma\mu_y B_0 - \frac{\mu_x}{T_2}\,, \\ \frac{\mathrm{d}\mu_y}{\mathrm{d}t} &= \gamma\mu_x B_0 - \frac{\mu_y}{T_2}\,, \\ \frac{\mathrm{d}\mu_z}{\mathrm{d}t} &= -\frac{\mu_z - \langle\mu_z\rangle}{T_1}\,. \end{aligned} \tag{3.2.111}$$

Their solution is

$$\begin{aligned} \mu_x &= A\cos\omega_{\mathrm{L}} t \cdot \mathrm{e}^{-t/T_2}\,, \\ \mu_y &= A\sin\omega_{\mathrm{L}} t \cdot \mathrm{e}^{-t/T_2}\,, \\ \mu_z &= \langle\mu_z\rangle\left[1 - B\mathrm{e}^{-t/T_1}\right]. \end{aligned} \tag{3.2.112}$$

The oscillation of angular frequency ω_{L} is indeed damped exponentially, and also the z component approaches its saturation value exponentially – however, the two relaxation times are different.

When precession is maintained via the application of a field of angular frequency ω, a forced oscillation is obtained. To simplify the solution of the

[21] F. Bloch, 1946. Felix Bloch (1905–1983) shared the Nobel Prize with Edward Mills Purcell (1912–1997) in 1952 "for their development of new methods for nuclear magnetic precision measurements and discoveries in connection therewith".

equations of motion, the linearly polarized field of frequency ω is replaced by two fields of amplitude B_1 that are circularly polarized and counterrotating in the (x, y) plane. Only the component corotating with the precession direction will have important effects, the other can be neglected. In the circularly polarized field obtained this way the equations of motion are

$$\begin{aligned}
\frac{\mathrm{d}\mu_x}{\mathrm{d}t} &= -\gamma\mu_y B_0 + \gamma\mu_z B_1 \sin\omega t - \frac{\mu_x}{T_2}\,, \\
\frac{\mathrm{d}\mu_y}{\mathrm{d}t} &= \gamma\mu_x B_0 - \gamma\mu_z B_1 \cos\omega t - \frac{\mu_y}{T_2}\,, \\
\frac{\mathrm{d}\mu_z}{\mathrm{d}t} &= \gamma\mu_y B_1 \cos\omega t - \gamma\mu_x B_1 \sin\omega t - \frac{\mu_z - \langle\mu_z\rangle}{T_1}\,.
\end{aligned} \tag{3.2.113}$$

As the magnetic moment corotates with the perturbing field, it is convenient to switch to a coordinate system rotating with angular frequency ω. The components in the two frames are related by

$$\mu_{x'} = \mu_x \cos\omega t + \mu_y \sin\omega t\,, \qquad \mu_{y'} = \mu_y \cos\omega t - \mu_x \sin\omega t \tag{3.2.114}$$

and

$$\mu_x = \mu_{x'} \cos\omega t - \mu_{y'} \sin\omega t\,, \qquad \mu_y = \mu_{y'} \cos\omega t + \mu_{x'} \sin\omega t\,. \tag{3.2.115}$$

Thus, in the rotating reference frame

$$\frac{\mathrm{d}'\boldsymbol{\mu}}{\mathrm{d}t} = \frac{\mathrm{d}\boldsymbol{\mu}}{\mathrm{d}t} - \boldsymbol{\omega} \times \boldsymbol{\mu}\,, \tag{3.2.116}$$

where $\boldsymbol{\omega}$ is a vector of magnitude ω along the constant magnetic field. In this reference frame the equations of motion are

$$\begin{aligned}
\frac{\mathrm{d}'\mu_{x'}}{\mathrm{d}t} &= -\gamma\mu_{y'} B_0 + \omega\mu_{y'} - \frac{\mu_{x'}}{T_2}\,, \\
\frac{\mathrm{d}'\mu_{y'}}{\mathrm{d}t} &= \gamma\mu_{x'} B_0 - \gamma\mu_z B_1 - \omega\mu_{x'} - \frac{\mu_{y'}}{T_2}\,, \\
\frac{\mathrm{d}'\mu_z}{\mathrm{d}t} &= \gamma\mu_{y'} B_1 - \frac{\mu_z - \langle\mu_z\rangle}{T_1}\,.
\end{aligned} \tag{3.2.117}$$

It is immediately seen that in the rotating coordinate system the moment precesses as if acted upon by an effective field $\boldsymbol{B}_{\mathrm{eff}} = \boldsymbol{B} - \boldsymbol{\omega}/\gamma$.

Assuming again that the temporal variations of the z component can be neglected, a stationary solution is found in the rotating reference frame:

$$\begin{aligned}
\mu_{x'} &= \frac{\gamma B_1 (\omega - \omega_{\mathrm{L}}) T_2^2}{1 + (\omega - \omega_{\mathrm{L}})^2 T_2^2 + \gamma^2 B_1^2 T_1 T_2} \langle\mu_z\rangle\,, \\
\mu_{y'} &= -\frac{\gamma B_1 T_2}{1 + (\omega - \omega_{\mathrm{L}})^2 T_2^2 + \gamma^2 B_1^2 T_1 T_2} \langle\mu_z\rangle\,, \\
\mu_z &= \frac{1 + (\omega - \omega_{\mathrm{L}})^2 T_2^2}{1 + (\omega - \omega_{\mathrm{L}})^2 T_2^2 + \gamma^2 B_1^2 T_1 T_2} \langle\mu_z\rangle\,.
\end{aligned} \tag{3.2.118}$$

Going back to the original coordinate system, we have

$$\mu_x = \frac{T_2^2}{1+(\omega-\omega_{\rm L})^2T_2^2+\gamma^2B_1^2T_1T_2}\left[(\omega-\omega_{\rm L})\cos\omega t+\frac{1}{T_2}\sin\omega t\right]\gamma B_1\langle\mu_z\rangle\,, \tag{3.2.119}$$

$$\mu_y = \frac{T_2^2}{1+(\omega-\omega_{\rm L})^2T_2^2+\gamma^2B_1^2T_1T_2}\left[(\omega-\omega_{\rm L})\sin\omega t-\frac{1}{T_2}\cos\omega t\right]\gamma B_1\langle\mu_z\rangle\,.$$

These yield

$$\begin{aligned}\chi'_{xx} &= \frac{\chi_0}{2}\omega_{\rm L}T_2\frac{(\omega_{\rm L}-\omega)T_2}{1+(\omega-\omega_{\rm L})^2T_2^2+\gamma^2B_1^2T_1T_2}\,,\\ \chi''_{xx} &= \frac{\chi_0}{2}\omega_{\rm L}T_2\frac{1}{1+(\omega-\omega_{\rm L})^2T_2^2+\gamma^2B_1^2T_1T_2}\,.\end{aligned} \tag{3.2.120}$$

Introducing the notation

$$\Gamma = \frac{2}{T_2}\sqrt{1+\gamma^2B_1^2T_1T_2}\approx\frac{2}{T_2}\,, \tag{3.2.121}$$

the real and imaginary parts of susceptibility are found to be

$$\begin{aligned}\chi'_{xx} &= \frac{\chi_0}{2}\omega_{\rm L}\frac{(\omega_{\rm L}-\omega)}{(\omega-\omega_{\rm L})^2+(\Gamma/2)^2}\,,\\ \chi''_{xx} &= \frac{\chi_0}{2}\omega_{\rm L}\frac{1}{T_2}\frac{1}{(\omega-\omega_{\rm L})^2+(\Gamma/2)^2}\,.\end{aligned} \tag{3.2.122}$$

Figure 3.2 shows the frequency dependence of the real and imaginary parts of susceptibility for typical values of the magnetic field and the relaxation time.

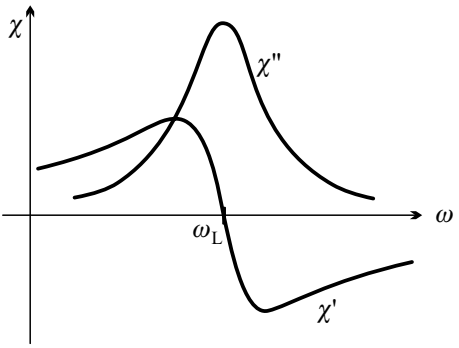


Fig. 3.2. Frequency dependence of the real and imaginary parts of susceptibility

Thus, instead of the infinitely sharp resonance obtained with the neglect of relaxation processes, a Lorentzian form is obtained for the imaginary part. In

weak perturbing fields the full width at half maximum (FWHM) is given by $1/T_2$. The line width is therefore a measure of the spin–spin relaxation time. In the relaxation-free limit $T_2 \to \infty$ ($\Gamma \to 0$) (3.2.108) is recovered for the imaginary part – however, the real part is now not precisely (3.2.102), as the field counterrotating with the magnetic moment has been neglected. The two expressions are, nonetheless, identical in the vicinity of the resonance (i.e., for frequencies $\omega \approx \omega_{\mathrm{L}}$) indicating that the approximation is justified.

The foregoing analysis showed that for a given external field there is a single resonant frequency. The selection rule $\Delta m_J = \pm 1$ implies that transition is possible only between adjacent levels. As the energy difference $\Delta E = g_J \mu_{\mathrm{B}} B$ is independent of m_J, the spectrum would consist of a single line for any value of J. In solids, however, this may change. The electrostatic potential due to neighboring atoms modifies the ground state of paramagnetic ions, and in particular, may partially lift the $2J + 1$-fold degeneracy. If the resulting energy shift is smaller than that caused by the magnetic field, absorption may occur at several nearby frequencies. This gives the fine structure of paramagnetic resonance. Atomic levels are further split by the hyperfine interaction with the nucleus of magnetic moment $\boldsymbol{\mu}_I$. This gives rise to the hyperfine structure of the absorption spectrum. The location of the resonance thus indicates the species of the paramagnetic ion as well as the properties of its surroundings, while its line width yields information about the interactions responsible for relaxation.

3.3 The Role of Nuclei

So far we have only discussed the states of core electrons. The nucleus has been assumed to be a pointlike and structureless object whose charge – together with that of the electrons – determines the total charge of the ion. However, the nucleus also possesses a magnetic moment through which it can interact with the electrons' spin. The role of the nucleus will also be important in the study of the structure and dynamics of solids by neutron scattering. Below we shall only consider the consequences of magnetic interactions.

3.3.1 Interaction with Nuclear Magnetic Moments

Both nucleon types, protons and neutrons, possess a magnetic moment. When expressed in terms of the nuclear magneton $\mu_{\mathrm{N}} = e\hbar/2m_{\mathrm{p}}$, the magnetic moment of the proton is $\mu_{\mathrm{p}} = 2.793\,\mu_{\mathrm{N}}$, while that of the neutron is $\mu_{\mathrm{n}} = -1.913\,\mu_{\mathrm{N}}$. Consequently most of the nuclei also possess a magnetic moment. The magnetic moment of a nucleus of angular momentum (nuclear spin) $\hbar\boldsymbol{I}$ is customarily written as

$$\boldsymbol{\mu}_I = g_I \mu_{\mathrm{N}} \boldsymbol{I}\,, \tag{3.3.1}$$

where g_I is the g-factor of the nucleus. Protons and neutrons are spin-1/2 particles, thus their g-factors are $g_{\mathrm{p}} = 5.586$ and $g_{\mathrm{n}} = -3.826$. For most

nuclei, the g-factor is of order unity, and the quantum number I can take integer or half-integer values.

Owing to the great disparity between the proton and electron mass, the nuclear magneton is three orders of magnitude smaller than the Bohr magneton, $\mu_{\mathrm{N}} \approx \mu_{\mathrm{B}}/1836$. For this reason one can usually ignore nuclear magnetic moments when studying magnetic properties. However, there are some situations where the magnetic field due to nuclear magnetic moments becomes important. As we have already mentioned, such fields give rise to the hyperfine structure of energy levels.[22] Using the Hamiltonian (3.2.19) of an electron in a magnetic field and the Coulomb gauge $\operatorname{div} \boldsymbol{A} = 0$,

$$\mathcal{H}_{\mathrm{hf}} = \frac{e}{m_{\mathrm{e}}}\, \boldsymbol{p} \cdot \boldsymbol{A}(\boldsymbol{r}) - \boldsymbol{B}(\boldsymbol{r}) \cdot \boldsymbol{\mu}_{\mathrm{s}}\,, \tag{3.3.2}$$

gives the nonrelativistic contribution to the hyperfine interaction that is proportional to the nuclear magnetic moment. The field $\boldsymbol{B}(\boldsymbol{r})$ due to a nucleus of magnetic moment $\boldsymbol{\mu}_I$ is derived from the vector potential

$$\boldsymbol{A}(\boldsymbol{r}) = \frac{\mu_0}{4\pi} \frac{\boldsymbol{\mu}_I \times \boldsymbol{r}}{r^3} = -\frac{\mu_0}{4\pi}\, \boldsymbol{\mu}_I \times \boldsymbol{\nabla} \frac{1}{r}\,. \tag{3.3.3}$$

Here $\boldsymbol{r}$ is the position vector relative to the nucleus. The first term in (3.3.2) represents the spin–orbit interaction between the nuclear spin and the orbiting electron, while the second term is the spin–spin (or dipole–dipole) coupling.

Substitution of the above expression for the vector potential into the Hamiltonian of the hyperfine interaction and rearrangement of the scalar triple product gives the following form for the first term:

$$\mathcal{H}^{(1)}_{\mathrm{hf}} = \frac{e\mu_0}{4\pi m_{\mathrm{e}} r^3}\, \boldsymbol{p} \cdot [\boldsymbol{\mu}_I \times \boldsymbol{r}] = \frac{e\mu_0 \hbar}{4\pi m_{\mathrm{e}} r^3}\, \boldsymbol{l} \cdot \boldsymbol{\mu}_I\,. \tag{3.3.4}$$

This expression can also be viewed as the energy of the nuclear magnetic moment $\boldsymbol{\mu}_I$ in the magnetic field

$$\boldsymbol{B}(\boldsymbol{r}) = \frac{\mu_0}{4\pi} \frac{-e\hbar \boldsymbol{l}}{m_{\mathrm{e}} r^3} \tag{3.3.5}$$

created by orbiting electron charges.

In the second term of (3.3.2) $\boldsymbol{B}(\boldsymbol{r})$ is derived from the previous expression for the vector potential as $\boldsymbol{B} = \boldsymbol{\nabla} \times \boldsymbol{A}$. Making use of the vector identity

$$\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) = (\boldsymbol{a} \cdot \boldsymbol{c})\boldsymbol{b} - (\boldsymbol{a} \cdot \boldsymbol{b})\boldsymbol{c} \tag{3.3.6}$$

that also holds when some of the vectors are the operator $\boldsymbol{\nabla}$, we have

[22] The electric quadrupole moment of nonspherically symmetric nuclei also contributes to the hyperfine structure – but here we are concerned with magnetic interactions alone.

$$\mathcal{H}_{\rm hf}^{(2)} = -\frac{\mu_0}{4\pi}\left\{(\boldsymbol{\mu}_I\cdot\boldsymbol{\nabla})(\boldsymbol{\mu}_{\rm s}\cdot\boldsymbol{\nabla}) - (\boldsymbol{\mu}_I\cdot\boldsymbol{\mu}_{\rm s})\boldsymbol{\nabla}^2\right\}\frac{1}{r}\,. \tag{3.3.7}$$

According to (C.3.8), $\boldsymbol{\nabla}^2(1/r)$ vanishes everywhere except at $r=0$, thus with the exception of this point

$$\mathcal{H}_{\rm hf}^{(2)} = -\frac{\mu_0}{4\pi}\left(\frac{3(\boldsymbol{\mu}_I\cdot\boldsymbol{r})(\boldsymbol{\mu}_{\rm s}\cdot\boldsymbol{r})}{r^5} - \frac{(\boldsymbol{\mu}_I\cdot\boldsymbol{\mu}_{\rm s})}{r^3}\right), \tag{3.3.8}$$

which is indeed a dipole–dipole interaction between the nuclear magnetic moment and the electron spin.

At $r=0$ the Laplacian is $\boldsymbol{\nabla}^2(1/r) = -4\pi\delta(\boldsymbol{r})$, thus the second term in $\mathcal{H}_{\rm hf}^{(2)}$ gives a singular contribution on the order of

$$-\frac{\mu_0}{4\pi}4\pi\boldsymbol{\mu}_I\cdot\boldsymbol{\mu}_{\rm s}\,\delta(\boldsymbol{r})\,. \tag{3.3.9}$$

Another singular contribution comes from the first term of (3.3.7). It can be evaluated using the Fourier representation of $1/r$, as given in (C.1.65):

$$\lim_{r\to0}(\boldsymbol{\mu}_I\cdot\boldsymbol{\nabla})(\boldsymbol{\mu}_{\rm s}\cdot\boldsymbol{\nabla})\frac{1}{r} = -\lim_{r\to0}\frac{1}{2\pi^2}\int\frac{{\rm e}^{{\rm i}\boldsymbol{q}\cdot\boldsymbol{r}}}{\boldsymbol{q}^2}(\boldsymbol{\mu}_I\cdot\boldsymbol{q})(\boldsymbol{\mu}_{\rm s}\cdot\boldsymbol{q})\,{\rm d}\boldsymbol{q}\,. \tag{3.3.10}$$

When the parentheses are expanded,

$$(\boldsymbol{\mu}_I\cdot\boldsymbol{q})(\boldsymbol{\mu}_{\rm s}\cdot\boldsymbol{q}) = \sum_{\alpha\beta}\mu_{I,\alpha}\mu_{{\rm s},\beta}q_\alpha q_\beta\,, \tag{3.3.11}$$

and the limit $r\to0$ is taken, only the terms $\alpha=\beta$ contribute to the integral on the right-hand side of (3.3.10):

$$\begin{aligned}\lim_{r\to0}\int{\rm e}^{{\rm i}\boldsymbol{q}\cdot\boldsymbol{r}}\frac{q_\alpha q_\beta}{\boldsymbol{q}^2}\mu_{I,\alpha}\mu_{{\rm s},\beta}\,{\rm d}\boldsymbol{q} &= \delta_{\alpha\beta}\frac{1}{3}\lim_{r\to0}\int{\rm e}^{{\rm i}\boldsymbol{q}\cdot\boldsymbol{r}}\mu_{I,\alpha}\mu_{{\rm s},\alpha}\,{\rm d}\boldsymbol{q}\\ &= \delta_{\alpha\beta}\frac{(2\pi)^3}{3}\mu_{I,\alpha}\mu_{{\rm s},\alpha}\delta(\boldsymbol{r})\,,\end{aligned} \tag{3.3.12}$$

hence

$$\lim_{r\to0}\left\{(\boldsymbol{\mu}_I\cdot\boldsymbol{\nabla})(\boldsymbol{\mu}_{\rm s}\cdot\boldsymbol{\nabla})\right\}\frac{1}{r} = -\frac{4\pi}{3}\delta(\boldsymbol{r})(\boldsymbol{\mu}_I\cdot\boldsymbol{\mu}_{\rm s})\,. \tag{3.3.13}$$

When all these contributions are collected, the interaction between the nuclear magnetic moment $\boldsymbol{\mu}_I$ and the electrons is written as

$$\begin{aligned}\mathcal{H}_{\rm hf} = -\frac{\mu_0}{4\pi}\Bigg\{&\frac{-e\hbar}{m_{\rm e}r^3}\boldsymbol{l}\cdot\boldsymbol{\mu}_I + \left[\frac{3(\boldsymbol{\mu}_{\rm s}\cdot\boldsymbol{r})(\boldsymbol{\mu}_I\cdot\boldsymbol{r})}{r^5} - \frac{(\boldsymbol{\mu}_{\rm s}\cdot\boldsymbol{\mu}_I)}{r^3}\right]\\ &+\frac{8\pi}{3}\boldsymbol{\mu}_{\rm s}\cdot\boldsymbol{\mu}_I\,\delta(\boldsymbol{r})\Bigg\}\,.\end{aligned} \tag{3.3.14}$$

The last term is the *Fermi contact term.*[23] It is nonzero only when the electron's wavefunction overlaps with that of the nucleus, therefore it is important only for s electrons.

3.3.2 Nuclear Magnetic Resonance

In perfect analogy to the behavior of the paramagnetic moment of the electron shell in a magnetic field, the nuclear magnetic moment also precesses around the direction of the applied magnetic field. If the nucleus is simultaneously subjected to an electromagnetic radiation of finite frequency, it can absorb energy as long as the frequency of the radiation is equal to the precession frequency. This phenomenon is called *nuclear magnetic resonance* (NMR).[24] The principle is the same as for electron spin resonance. The only apparent difference is that the resonance frequency is determined from the expression

$$\hbar\omega = g_J \mu_\mathrm{B} B \tag{3.3.15}$$

for ESR and from

$$\hbar\omega = g_I \mu_\mathrm{N} B \tag{3.3.16}$$

for NMR. In a magnetic field of 1 tesla, the resonant frequency for a free electron spin is $\nu_\mathrm{L} = 28\,\mathrm{GHz}$. Because of the huge disparity between the Bohr magneton and the nuclear magneton, the resonance for protons occurs at a much lower frequency, 42.6 MHz. This means that for field strengths commonly used in experiments – B is typically a few (at most 10 to 15) teslas – measurements can be made in the radiofrequency (rather than the microwave) region.

The same Bloch equations are used in both cases, however, the physical processes that determine the relaxation times T_1 and T_2 are, naturally, completely different for the relaxation of the nuclear spin and for the relaxation of the paramagnetic moment of an electron shell. Instead of the spin–lattice interaction, the relaxation time T_1 is basically determined by the hyperfine interaction in metals. As it was shown by J. Korringa (1950), T_1 is then inversely proportional to temperature. This relation is called the *Korringa law*, and the phenomenon itself the *Korringa relaxation.*

Just as the electron spin feels the presence of the nuclear spin, the field felt by the nuclear spin also depends on the polarization of the electron states. This will be of particular interest in metals where s-electrons have a nonvanishing density at the nucleus, therefore their polarization induced by the uniform

[23] E. Fermi, 1930. Enrico Fermi (1901–1954) was awarded the Nobel Prize in 1938 "for his demonstrations of the existence of new radioactive elements produced by neutron irradiation, and for his related discovery of nuclear reactions brought about by slow neutrons".

[24] See footnote on page 65.

field gives rise to an additional energy shift through the Fermi contact term. This leads to the *Knight shift*[25] of the NMR line in metals.

NMR does not have the same scope of applications as ESR since in solid-state physics (and in many applications) one is not interested in the magnetic moment of the nucleus but rather in the internal, local magnetic field within the sample, which can be inferred from the location of the resonance. Compared to a free atom, this frequency is shifted as neighboring para- or diamagnetic atoms/ions cause tiny variations in the effective magnetic field felt by the nucleus. This *chemical shift* being characteristic of the local environment of the nucleus, its measurement can yield information about it. Nuclear magnetic resonance methods are therefore extensively used in chemical, biological, and medical applications.

Further information can be gained about the local environment if the nucleus also possesses an electric quadrupole moment, since the latter interacts with the gradient of the crystalline electric field at the nucleus. This interaction modifies the location and width of the absorption peak measured by NMR.

3.3.3 The Mössbauer effect

As we have already mentioned, the most important field of application of NMR is the mapping of the vicinity of a nucleus within a bulk matter (in many cases a living tissue). Another method for studying the local environment is provided by the Mössbauer effect[26] (1957) through its influence on the transitions between the energy levels of the nucleus. In contrast to NMR, here one is concerned with transitions between states with different quantum numbers I – and not between the sublevels of different magnetic quantum numbers m_I into which the nuclear level I is split in a magnetic field. The most commonly used isotope is iron-57 (^{57}Fe) – however the phenomenon can be studied on a handful of other isotopes as well.

Upon the capture of an electron, the radioactive ^{57}Co nucleus decays into a spin-5/2 excited state of the ^{57}Fe isotope, which, in turn, quickly decays into a spin-3/2 excited state through the emission of a high-energy γ photon. The $I = 3/2$ level is somewhat longer lived ($\tau = 10^{-7}$ s), and will eventually make a transition to the $I = 1/2$ ground state through the emission of a 14.4 keV γ photon.

The process in which a photon absorbed by an atom is re-emitted at the same frequency is called *resonance fluorescence*. The inverse process, in which a photon emitted by an atom is absorbed by an identical atom, is called *resonance absorption*. Such processes occur with high probability when the

[25] W. D. Knight, 1949.

[26] Rudolf Ludwig Mössbauer (1929–) was awarded the Nobel Prize in 1961 "for his researches concerning the resonance absorption of gamma radiation and his discovery in this connection of the effect which bears his name".

photon is emitted and absorbed in a transition in the electron shell. Although momentum is conserved in emission and absorption processes alike, and thus the atoms will be recoiled, the recoil energy is nonetheless smaller than the natural line width. The situation is different for γ rays emitted by individual nuclei. Here the recoil energy can be several orders of magnitude larger than the natural line width – for the above mentioned transition of ^{57}Fe, $\Gamma = 10^{-8}\,\mathrm{eV}$, while the recoil energy is $2\,\mathrm{meV}$ –, and so the photon emitted by a free nucleus could not be absorbed by another free identical nucleus in its ground state. As MÖSSBAUER pointed out, the situation is radically different when the emitting and absorbing isotopes are embedded in crystal lattices. In this case the momentum is taken up by the entire lattice, making the recoil energy negligible. Thus at low temperatures, where the lattice vibrations are small, resonance absorption may occur with high probability. This is why the phenomenon is also referred to as the recoilless emission and nuclear resonance absorption of gamma radiation.

To be more precise, there can be a small difference between the frequencies of the emitted and absorbable photons. Since the size of the nucleus is slightly but distinctly different in the $I = 3/2$ and $I = 1/2$ states, the energy corrections due to the Coulomb interaction between the nucleus and the electron cloud are different in the excited state and in the ground state. The energy difference between the two states is proportional to the difference of the nuclear radii, and also depends on the electron density at the nucleus. If the electron densities at the emitter and absorber nuclei are different, *isomer shift* – the frequency difference of the emitted and absorbable rays – can prevent absorption. This can be remedied by moving the radiation source relative to the sample: the Doppler shift caused by the relative velocity of the emitter and the absorber can compensate for the isomer shift.

If a local magnetic field is present at the nucleus, it will split the energy levels of both the ground and excited states by an amount called the Zeeman energy. Moreover, if the electric field has a finite gradient at the nucleus because of the crystalline surroundings, then – due to the quadrupole moment of the $I = 3/2$ state – the excited levels $m_I = \pm 3/2$ and $m_I = \pm 1/2$ will suffer additional shifts in opposite directions. The arising level structure is shown schematically in Fig. 3.3(a). Due attention has been paid to the sign of the magnetic moment: positive in the ground state ($\mu = 0.09\,\mu_\mathrm{N}$) and negative in the excited state ($\mu = -0.15\,\mu_\mathrm{N}$). The selection rules $\Delta m_I = 0, \pm 1$ allow six transitions among these levels. The transition energies are shifted by different amounts with respect to the transition energy in a free atom but the Doppler effect can compensate for these shifts as well. From measurements of the extent of the splitting and of the relative intensities of these lines one can determine the internal local fields, their direction and temperature dependence, and from these specify the local environment. As an example, consider Fig. 3.3(b) showing the Mössbauer spectrum of an iron sample containing ^{57}Fe, measured at various temperatures using a ^{57}Co source. Above the Curie temperature, where no internal magnetic field is present, a single line appears.

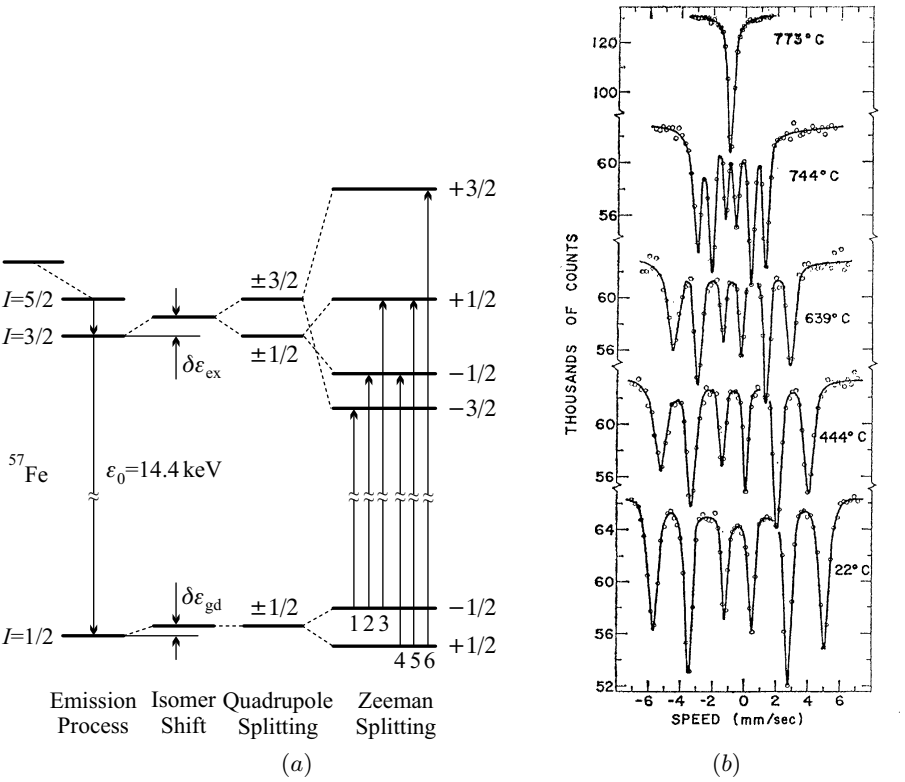


Fig. 3.3. (*a*) Splitting of the levels of ^{57}Fe in the presence of a local magnetic field and a finite-gradient electric field. (*b*) Typical Mössbauer spectra for metallic ^{57}Fe between room temperature and the Curie temperature [R. S. Preston, S. S. Hanna, and J. Heberle, *Phys. Rev.* **128**, 2207 (1962)]

At lower temperatures the sample becomes spontaneously magnetized. The internal magnetic field splits the nuclear levels, and six distinct lines appear; they correspond to the six allowed transitions.

Further Reading

1. C. P. Slichter, *Principles of Magnetic Resonance*, Third Enlarged and Updated Edition, Corrected 3rd printing, Springer Series in Solid-State Sciences, Vol. 1, Springer-Verlag, Berlin (1996).
2. K. Yosida, *Theory of Magnetism*, corrected 2nd printing, Springer Series in Solid-State Sciences, Springer-Verlag, Berlin (1998).
3. J. H. Van Vleck, *The Theory of Electric and Magnetic Susceptibilities*, Clarendon Press, Oxford (1932).

4

Bonding in Solids

The foregoing discussion of the electron shells of individual atoms and ions belongs essentially to the realm of atomic or molecular physics. Condensed matter physics is principally concerned with what happens or changes when atoms are placed in the vicinity of each other, at distances characteristic of solids. Experience shows that for such densities direct interactions between the ion cores and effective ones mediated by unbound electrons create stable, rigid structures. In this chapter we shall give a brief overview of the types of bonds electrons can create between atoms, ions, or molecules, and then discuss the methods used for calculating the cohesive energy – the measure of stability in solids. We shall return to the relation between the dominant type of bonding in the solid and its structure after the presentation of crystalline structures, in Chapter 7.

4.1 Types of Bonds and Cohesive Energy

One can imagine building up solids from a free gas of atoms or molecules by letting them approach each other until they are at the positions they would occupy within the solid. During this procedure the state of electrons on closed inner shells does not change considerably. On the other hand, the state of electrons on incompletely filled outer shells may change essentially. Such electrons may be stripped off from the atomic core leaving behind a positive ion, and become more or less free, *delocalized*. They may equally become localized around another atom, thereby changing its charge. In other cases they may create bonds between two atoms that are similar to covalent chemical bonds in molecules – then the density of the electrons involved in the bond (the bonding electrons) has its maximum in the region between the two atoms. In each case, the bonding of atomic cores is ensured by electrons. One may say that electrons are the glue that holds ion cores together via Coulomb interactions.

Depending on the character of the electron states within the solid that are primarily responsible for bonding, the following types of bonds are distinguished: (i) van der Waals bond, (ii) ionic bond, (iii) covalent bond, (iv) metallic bond.

For completeness, at the end of the chapter we shall briefly discuss a particular possibility for chemical bonding, the so-called hydrogen bond. This plays an important role in biological substances but hardly any in solid-state physics. Bonding between atomic cores is then ensured by hydrogen bonds created by H^+ ions (bare protons), and not by electrons.

4.1.1 Classification of Solids According to the Type of the Bond

The properties of solids are determined to a large extent by the bonds that hold their constituents together. Thus the classification of bonds gives a possible classification of solids as well. However, it should be borne in mind that in many cases bonds do not appear in a clear-cut manner. Ionic and covalent bonds can be regarded as two extremes between which a continuity exists. It is not uncommon either that different types of bonding occur along different directions in the same sample. For example, in graphite covalent bonds are found within the planes but only van der Waals forces between layers. According to the dominant type of bonding, the following classes are distinguished: (i) molecular crystals, (ii) ionic crystals, (iii) covalent crystals, (iv) metals.

In molecular crystals closed-shell atoms or molecules are held together by van der Waals forces. Noble gases form such molecular crystals in their solid state – and so do certain large molecules, e.g. soccer-ball-shaped fullerene molecules (C_{60}) (see page 30).

The other names are self-explanatory. In ionic crystals the dominant bonding is ionic. The most prominent representatives of this class are alkali halides. Donor–acceptor salts are also held together by ionic bonds; in this case the regular crystalline structure is formed by donor molecules that easily give away electrons and acceptor molecules that easily take them up. Special mention should be made of TTF-TCNQ and its derivatives: their particular properties – to be discussed in detail in Chapter 33 – opened up the door through which organic materials found their way into solid-state physics.

A crystal is called covalent if the ion cores are bound together by covalent chemical bonds formed by valence electrons localized along the bond. Owing to the saturated character of covalent bonds such materials are insulators. When bonds are unsaturated the material is a conductor. However, such a simple explanation cannot account for the forces that hold metals together. That is why metallic bond forms a separate category, in which delocalized electrons play the primary role.

4.1.2 Cohesive Energy

Solids are usually thermally stable up to a certain temperature. The attractive interaction among atoms gives rise to a lower energy in a solid than in a

liquid, which is less bonded, or in a gas, which is made up of practically freely moving atoms or molecules. The extent by which the energy of the solid state is lower than the energy of the set of independent constituent atoms is called the *cohesive energy*, since it accounts for the cohesion of the solid. In the literature it is customarily given as the energy (usually in units of eV) required to remove one atom or molecule – or, alternatively, as the cohesive energy for one mole of substance (in units of kJ/mol). For convenience, we shall use both definitions below. The conversion factor is[1]

$$1\,\frac{\text{eV}}{\text{atom}} = 96.4853\,\frac{\text{kJ}}{\text{mol}}\,. \tag{4.1.1}$$

A precise calculation of the cohesive energy is rather difficult. Besides the energies of the bonds formed by electrons between the atoms, due account must be taken of the oscillatory motion of the atoms within the crystal, which also contributes to the total energy of the solid. Its measurement is not straightforward either, therefore data cited by different sources may significantly differ.

Table 4.1. Cohesive energy of some chemical elements (in units of eV/atom)

Li	Be									C	O	F	Ne		
1.66	3.33									7.35	2.60	0.84	0.020		
Na	Mg									Si	S	Cl	Ar		
1.13	1.53									4.66	2.86	1.40	0.080		
K	Ca	Fe	Co	Ni	Cu					Ge	Se	Br	Kr		
0.94	1.83	4.31	4.41	4.44	3.50					3.85	2.13	1.22	0.116		
Rb	Sr	Ru	Rh	Pd	Ag					Sn	Te	I	Xe		
0.88	1.70	6.68	5.77	3.90	2.96					3.12	2.02	1.11	0.17		
Cs	Ba	Os	Ir	Pt	Au					Pb	Po	At	Rn		
0.83	1.86	8.12	6.89	5.85	3.80					2.04	1.50	N/A	0.202		

Table 4.1 clearly shows that cohesive energies are substantially different for "noble gases" – that form molecular crystals in the solid state –, for simple or transition metals, and covalently bonded elemental semiconductors. The cohesive energy of noble gases is the smallest of all, on the order of 0.1 eV per atom. For simple metals it is around 1 eV per atom, and even larger for transition metals and covalently bonded crystals.

Similar or even larger cohesive energies are found in covalently bonded crystals with a partially ionic character, and also in crystals with purely ionic bonds. For the latter, it is often not the cohesive energy that is listed in tables but the *separation energy*, which is the energy difference between the total crystal energy per molecule and the energy of a free molecule. Table 4.2 contains the cohesive and separation energies for some purely ionic alkali halide crystals and some compounds with partially ionic bonds.

[1] Units of kcal/mol are often used in older tables. This can be converted to units of kJ/mol using the conversion factor 1 kcal = 4.184 kJ.

Table 4.2. Low-temperature cohesive and separation energies for some purely ionic alkali halide crystals and some compounds with partially ionic bonds (in units of kJ/mol and eV/molecule)

Substance	Cohesive energy		Substance	Cohesive energy	
	kJ/mol	eV/molecule		kJ/mol	eV/molecule
LiF	1036	10.74	LiCl	853	8.84
NaF	923	9.57	NaCl	786	8.15
KF	821	8.51	KCl	715	7.41
RbF	785	8.14	RbCl	689	7.14
CsF	740	7.67	CsCl	659	6.83
Substance	Separation energy		Substance	Separation energy	
	kJ/mol	eV/molecule		kJ/mol	eV/molecule
CaF_2	625	6.48	MgO	950	9.85
$PbCl_2$	521	5.40	Cu_2O	788	8.17
ZnS	852	8.83	Al_2O_3	2790	28.92

In the following sections we shall briefly discuss the various types of bonding, the fundamentals of the underlying quantum mechanics – without getting involved in the subtleties of quantum chemistry –, the theoretical methods for calculating the cohesive energy, and its expected order of magnitude.

4.2 Molecular crystals

As it has been already mentioned, stable and inert closed-shell molecules are held together by rather weak forces in molecular crystals. This weak interaction is a consequence of quantum fluctuations. A fluctuating dipole moment on one atom will polarize the closed electron shell of another, giving rise to an induced dipole moment. As the electric field of a dipole drops off as the inverse cube of the distance, the energy of attraction between a randomly appearing electric dipole and the dipole it induces is inversely proportional to the sixth power of their separation R:

$$\varepsilon = -\frac{C}{R^6}\,. \tag{4.2.1}$$

This interaction is called the *van der Waals interaction*[2] although its quantum mechanical interpretation is due to F. London (1930).

[2] J. D. van der Waals, 1873. Johannes Diderik van der Waals (1837–1923) was awarded the Nobel Prize in 1910 "for his work on the equation of state for gases and liquids".

4.2.1 Van der Waals Bonds in Quantum Mechanics

For a more precise description of the van der Waals interaction consider two noble-gas atoms, at points $\boldsymbol{R}_\mathrm{A}$ and $\boldsymbol{R}_\mathrm{B}$, separated by a distance $R = |\boldsymbol{R}| = |\boldsymbol{R}_\mathrm{B} - \boldsymbol{R}_\mathrm{A}|$. As we are concerned with the state of electrons only, the positions of atoms are assumed to be rigidly fixed and thus their kinetic energy is ignored throughout this chapter. As long as the size of the electron clouds are smaller than the separation R of the atoms, each electron can be unambiguously associated with one atom. The position vectors of the Z electrons that belong to the atom at $\boldsymbol{R}_\mathrm{A}$ ($\boldsymbol{R}_\mathrm{B}$) will be denoted by $\boldsymbol{r}_i^{(\mathrm{A})}$ ($\boldsymbol{r}_i^{(\mathrm{B})}$). Thus the Hamiltonian of the system is composed of three parts:

$$\mathcal{H} = \mathcal{H}_\mathrm{A} + \mathcal{H}_\mathrm{B} + \mathcal{H}_\mathrm{int}\,, \tag{4.2.2}$$

where $\mathcal{H}_\mathrm{A}$ contains the kinetic energy operator of the electrons that belong to the atom at $\boldsymbol{R}_\mathrm{A}$ plus the Coulomb interactions of these electrons with each other and with the nucleus:

$$\mathcal{H}_\mathrm{A} = -\sum_{i=1}^{Z} \frac{\hbar^2}{2m_\mathrm{e}} \boldsymbol{\nabla}_i^2 + \frac{1}{2} \sum_{\substack{i,j=1 \\ i \neq j}}^{Z} \frac{\tilde{e}^2}{\left|\boldsymbol{r}_i^{(\mathrm{A})} - \boldsymbol{r}_j^{(\mathrm{A})}\right|} - \sum_{i=1}^{Z} \frac{Z\tilde{e}^2}{\left|\boldsymbol{r}_i^{(\mathrm{A})} - \boldsymbol{R}_\mathrm{A}\right|}\,. \tag{4.2.3}$$

The Hamiltonian of the other atom, $\mathcal{H}_\mathrm{B}$ takes the same form, while $\mathcal{H}_\mathrm{int}$ contains the pairwise Coulomb interactions between the two nuclei, between the nuclei and the electrons bound to the other atom, and between the electrons bound to different atoms:

$$\mathcal{H}_\mathrm{int} = \frac{Z^2\tilde{e}^2}{|\boldsymbol{R}|} - \sum_{i=1}^{Z} \left(\frac{Z\tilde{e}^2}{\left|\boldsymbol{r}_i^{(\mathrm{A})} - \boldsymbol{R}_\mathrm{B}\right|} + \frac{Z\tilde{e}^2}{\left|\boldsymbol{r}_i^{(\mathrm{B})} - \boldsymbol{R}_\mathrm{A}\right|} \right) + \sum_{i,j=1}^{Z} \frac{\tilde{e}^2}{\left|\boldsymbol{r}_i^{(\mathrm{A})} - \boldsymbol{r}_j^{(\mathrm{B})}\right|}\,. \tag{4.2.4}$$

When the notations $\widetilde{\boldsymbol{r}}_i^{(\mathrm{A})} = \boldsymbol{r}_i^{(\mathrm{A})} - \boldsymbol{R}_\mathrm{A}$ and $\widetilde{\boldsymbol{r}}_i^{(\mathrm{B})} = \boldsymbol{r}_i^{(\mathrm{B})} - \boldsymbol{R}_\mathrm{B}$ are introduced, i.e., when the position vectors of the electrons are referred to the equilibrium position of their respective atoms, we have

$$\mathcal{H}_\mathrm{int} = \frac{Z^2\tilde{e}^2}{|\boldsymbol{R}|} - \sum_{i=1}^{Z} \left(\frac{Z\tilde{e}^2}{\left|\widetilde{\boldsymbol{r}}_i^{(\mathrm{A})} - \boldsymbol{R}\right|} + \frac{Z\tilde{e}^2}{\left|\widetilde{\boldsymbol{r}}_i^{(\mathrm{B})} + \boldsymbol{R}\right|} \right) + \sum_{i,j=1}^{Z} \frac{\tilde{e}^2}{\left|\widetilde{\boldsymbol{r}}_i^{(\mathrm{A})} - \widetilde{\boldsymbol{r}}_j^{(\mathrm{B})} - \boldsymbol{R}\right|}\,. \tag{4.2.5}$$

As R is assumed to be much larger than $|\widetilde{\boldsymbol{r}}_i^{(\mathrm{A})}|$ and $|\widetilde{\boldsymbol{r}}_i^{(\mathrm{B})}|$, we shall expand the previous expression into a power series in the electron position vectors. Using the expansions

$$\frac{1}{\left|\boldsymbol{R} - \widetilde{\boldsymbol{r}}_i^{(\mathrm{A})}\right|} = \frac{1}{R} - \sum_{\alpha} \widetilde{r}_{i\alpha}^{(\mathrm{A})} \frac{\partial}{\partial R_\alpha} \frac{1}{R} + \frac{1}{2} \sum_{\alpha\beta} \widetilde{r}_{i\alpha}^{(\mathrm{A})} \widetilde{r}_{i\beta}^{(\mathrm{A})} \frac{\partial^2}{\partial R_\alpha \partial R_\beta} \frac{1}{R} + \ldots \tag{4.2.6}$$

and

$$\frac{1}{\left|\boldsymbol{R}-\left(\widetilde{\boldsymbol{r}}_i^{(\mathrm{A})}-\widetilde{\boldsymbol{r}}_j^{(\mathrm{B})}\right)\right|}=\frac{1}{R}-\sum_{\alpha}\left(\widetilde{r}_{i\alpha}^{(\mathrm{A})}-\widetilde{r}_{j\alpha}^{(\mathrm{B})}\right)\frac{\partial}{\partial R_\alpha}\frac{1}{R} \tag{4.2.7}$$

$$+\frac{1}{2}\sum_{\alpha\beta}\left(\widetilde{r}_{i\alpha}^{(\mathrm{A})}-\widetilde{r}_{j\alpha}^{(\mathrm{B})}\right)\left(\widetilde{r}_{i\beta}^{(\mathrm{A})}-\widetilde{r}_{j\beta}^{(\mathrm{B})}\right)\frac{\partial^2}{\partial R_\alpha\partial R_\beta}\frac{1}{R}+\dots,$$

as well as the relations

$$\frac{\partial}{\partial R_\alpha}\frac{1}{R}=-\frac{n_\alpha}{R^2},\qquad \frac{\partial^2}{\partial R_\alpha\partial R_\beta}\frac{1}{R}=\frac{1}{R^3}\left[3n_\alpha n_\beta-\delta_{\alpha\beta}\right], \tag{4.2.8}$$

where $n_\alpha = R_\alpha/R$ is the α component of the unit vector $\boldsymbol{n}$ along the direction $\boldsymbol{R}$, it is easily seen that the zeroth- as well as the first-order terms cancel out, and the surviving second-order term contains the coordinates of the electrons of both atoms:

$$\mathcal{H}_{\mathrm{int}}=-\frac{\tilde{e}^2}{R^3}\sum_{ij}\left[3(\widetilde{\boldsymbol{r}}_i^{(\mathrm{A})}\cdot\boldsymbol{n})(\widetilde{\boldsymbol{r}}_j^{(\mathrm{B})}\cdot\boldsymbol{n})-\widetilde{\boldsymbol{r}}_i^{(\mathrm{A})}\cdot\widetilde{\boldsymbol{r}}_j^{(\mathrm{B})}\right]. \tag{4.2.9}$$

Introducing the dipole moments of the atoms by

$$\boldsymbol{d}^{(\mathrm{A})}=-e\sum_{i=1}^{Z}\widetilde{\boldsymbol{r}}_i^{(\mathrm{A})},\qquad \boldsymbol{d}^{(\mathrm{B})}=-e\sum_{i=1}^{Z}\widetilde{\boldsymbol{r}}_i^{(\mathrm{B})}, \tag{4.2.10}$$

the Hamiltonian of the interaction takes the usual form of dipole–dipole interactions,

$$\mathcal{H}_{\mathrm{int}}=-\frac{1}{R^3}\left[3(\boldsymbol{d}^{(\mathrm{A})}\cdot\boldsymbol{n})(\boldsymbol{d}^{(\mathrm{B})}\cdot\boldsymbol{n})-\boldsymbol{d}^{(\mathrm{A})}\cdot\boldsymbol{d}^{(\mathrm{B})}\right]. \tag{4.2.11}$$

This shows that the leading term in the interaction of two distant neutral atoms is the dipole–dipole interaction. For closed-shell atoms the dipole moment vanishes as their charge distribution is spherically symmetric. The first nonvanishing contribution to the energy appears in the second order of perturbation theory. This correction can be viewed as the interaction between the induced dipole moments of the electron clouds. A well-known theorem of quantum mechanical perturbation theory asserts that the second-order energy correction is always negative for ground-state atoms – which means that in this particular approximation the two atoms attract one another. The energy correction depends on separation as $1/R^6$, and its strength is determined by the induced dipole moments of the atoms.

An estimation of the matrix elements that appear in perturbation theory gives the interaction energy

$$E=-6\frac{\tilde{e}^2}{R}\left(\frac{a_0}{R}\right)^5, \tag{4.2.12}$$

where a_0 is the Bohr radius. As a_0 is smaller than usual interatomic distances, this energy is much smaller than usual Coulomb energies of order $\tilde{e}^2/R$. Weak

and short as it may be, the van der Waals force is the only one that can hold together neutral atoms or molecules with zero permanent dipole moment in a solid when interatomic distances are larger than the dimensions of atomic electron clouds (i.e., when each electron is essentially localized around one particular nucleus). Van der Waals forces act under other circumstances as well. However, their effects can be safely ignored when other interactions are also present.

4.2.2 Cohesive Energy of Molecular Crystals

If attractive $1/r^6$ interactions were present in molecular crystals all the way down to arbitrarily small scales, these crystals would collapse. This is obviously not the case. The derivation of the van der Waals interaction was based on the assumption that the electron clouds of the two molecules do not overlap. At small distances, when the electron shells would overlap, the molecules repel each other much like rigid spheres. In molecular crystals this short-range repulsion is customarily taken into account as an empirical term that drops off as a high power of distance. For computational ease – although not implied by any rigorous derivation – the most common choice is the *Lennard-Jones (6–12) potential*[3] that contains the van der Waals interaction and the short-distance repulsion in the form

$$U(r) = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^{6} \right]. \tag{4.2.13}$$

The spatial variation of the potential is shown in Fig. 4.1. Although there is no compelling reason for the choice of 12 as the power in the repulsive term, it provides a remarkably good description for the solid state of noble gases.

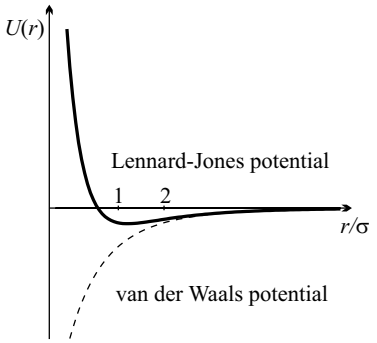


Fig. 4.1. The Lennard-Jones potential as a function of distance

The parameters ϵ and σ in the potential can be related to the virial coefficients of statistical physics, thus they can be determined from the behavior

[3] J. E. Lennard-Jones, 1924.

of the gas phase. The values thus obtained for these parameters are given in Table 4.3. When these are known, the properties of the solid phase can be determined without further fitting parameters.

Table 4.3. Values of the Lennard-Jones parameters ϵ and σ determined from the behavior of the gas state in noble gases. r_0 ($r_0^{\exp}$) and ε_0 ($\varepsilon_0^{\exp}$) are the calculated (measured) values of the smallest interatomic distance and the cohesive energy per atom, respectively

	Ne	Ar	Kr	Xe
ϵ (eV)	0.003	0.010	0.014	0.020
σ (Å)	2.78	3.40	3.60	4.10
r_0 (Å)	3.03	3.71	3.93	4.47
$r_0^{\exp}$ (Å)	3.155	3.756	3.992	4.336
ε_0 (eV)	−0.026	−0.086	−0.121	−0.172
$\varepsilon_0^{\exp}$ (eV)	−0.020	−0.080	−0.116	−0.17

To calculate the total energy of a molecular crystal, the potential energies felt by individual molecules have to be summed. Using the notation r_0 for the nearest-neighbor distance, the separation r_{ij} between more distant molecules is customarily written as $r_{ij} = p_{ij} r_0$. The total energy is then

$$\begin{aligned} E_0 &= \sum_{i \neq j} 4\epsilon \left[\left(\frac{\sigma}{p_{ij} r_0} \right)^{12} - \left(\frac{\sigma}{p_{ij} r_0} \right)^{6} \right] \\ &= \tfrac{1}{2} N 4\epsilon \left[A_{12} \left(\frac{\sigma}{r_0} \right)^{12} - A_6 \left(\frac{\sigma}{r_0} \right)^{6} \right], \end{aligned} \tag{4.2.14}$$

where the structure-dependent coefficients

$$A_n = \sum_j \left(\frac{1}{p_{ij}} \right)^n \tag{4.2.15}$$

are easily calculated, as the series converge rapidly because of the large exponents. For face-centered cubic crystals, where each atom has 12 nearest neighbors, $A_6 = 14.4539$ and $A_{12} = 12.1319$. For body-centered cubic structures, where the coordination number is 8, $A_6 = 12.2533$ and $A_{12} = 9.1142$.

The equilibrium distance between nearest neighbors obtained by minimizing the energy is

$$r_0 = \left(\frac{2A_{12}}{A_6} \right)^{1/6} \sigma \,. \tag{4.2.16}$$

With this value for the equilibrium lattice constant the energy per atom is

$$\varepsilon_0 = -\frac{A_6^2}{2A_{12}}\,\epsilon\,. \tag{4.2.17}$$

The above values for the coefficients A_6 and A_{12} give $r_0 = 1.09\,\sigma$ for face-centered cubic crystals. Using the σ obtained from the virial coefficients (which are, in turn, determined from the properties of the gas phase), (4.2.16) gives a good estimate for the lattice constants of noble-gas crystals. Experimental values for the energy per atom – listed in Table 4.3 – are also in good agreement with calculated ones, which are obtained from the parameter ϵ through $\varepsilon_0 = -8.61\,\epsilon$. Given the energy, elastic properties such as the bulk modulus can be determined, once again in good agreement with experimental data.

Owing to the weakness of van der Waals forces the cohesive energy per molecule is typically 0.1 eV in other molecular crystals, too. Therefore the melting point of molecular crystals is low.

4.3 Ionic Bond

Another type of bond is formed between two initially neutral atoms if their electronegativities are very different. (We shall refer to atoms but our considerations are equally valid for a pair of molecules, when one attracts electrons much more strongly than the other.) When the difference in electronegativity is larger than 1.7 on the Pauling scale,[4] the atom with higher electronegativity can attract to itself the relatively weakly bound electrons on the outermost partially filled electron shell of the second atom. This charge transfer process may continue up to the point when no partially filled shells are found on either ion. The crystal is held together by the Coulomb interaction between positive *cations* or negative *anions*. This is called *ionic bonding*.[5]

The most typical examples are alkali halide compounds formed by elements of group 1 (IA, alkali metals) and group 17 (VIIA, halogens) of the periodic table. The difference in electronegativity is smaller between the elements of groups 12 (IIB) and 16 (VIA), nonetheless it is still large enough for the complete transfer of two electrons. In other cases the outermost shells of the ions are not necessarily complete but the electrons are relatively well localized around the ions produced by the charge transfer. Such systems also belong to the same type of bonding provided the dominant force is the Coulomb interaction between oppositely charged ions.

In ionically bonded crystals ions form a lattice in which each cation (anion) is surrounded by as many anions (cations) as possible, in a symmetric arrangement. It is thus not possible to distinguish electrically neutral molecules

[4] L. C. Pauling, 1932. Linus Carl Pauling (1901–1994) was awarded the Nobel Prize in Chemistry in 1954 "for his research into the nature of the chemical bond and its application to the elucidation of the structure of complex substances". Eight years later he also received the Nobel Peace Prize.

[5] The alternative term, *heteropolar bonding* refers to the opposite charge of the ions.

within the lattice. The most important contribution to the cohesive energy of crystals is due to the Coulomb interactions between ions. Because of their long range, one must take into account the attraction and repulsion between arbitrarily distant ions when calculating the energy of the system. When site $\boldsymbol{R}_i$ is occupied by an ion of charge Z_i, the Coulomb interaction energy

$$E = \frac{1}{8\pi\epsilon_0} \sum_{i \neq j} \frac{Z_i Z_j e^2}{|\boldsymbol{R}_i - \boldsymbol{R}_j|} = \frac{1}{2} \sum_{i \neq j} \frac{Z_i Z_j \tilde{e}^2}{|\boldsymbol{R}_i - \boldsymbol{R}_j|} \tag{4.3.19}$$

is called the *Madelung energy* of the ionic crystal.[6] For most types of ionic crystals this is an asymptotic series in which slowly decreasing positive and negative terms alternate, due to the signs of Z_i and Z_j and the $1/r$ decay of the Coulomb potential. Although strictly speaking asymptotic series do not converge, fairly rapidly converging methods have been worked out for calculating of the sum of the series – i.e., the energy – by P. P. Ewald (1921) and H. M. Evjen (1932).

By way of example, consider two typical ionic crystals, NaCl and CsCl. In the sodium chloride (or rock-salt) crystal the sites of a regular cubic lattice are occupied by Na^+ and Cl^- ions alternately, in a 3D checkerboard pattern. In the CsCl crystal, Cs^+ ions make up a regular cubic lattice, and Cl^- ions occupy the center of each elementary cube. Figure 4.2 shows small portions of the two lattice types, as well as the distances between first- (nearest-), second-, third-, fourth-, and fifth-neighbor ions.

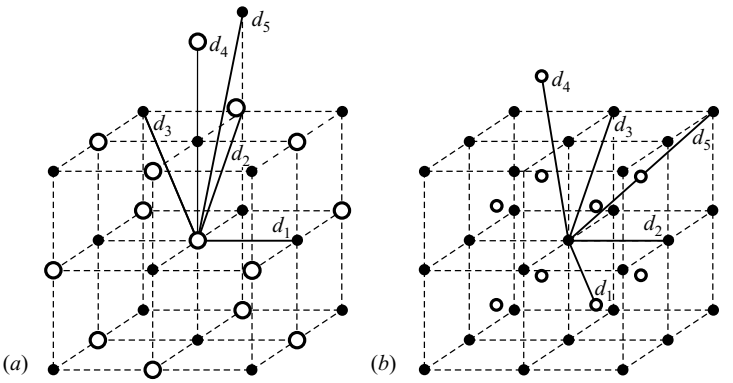


Fig. 4.2. The structure of two typical ionic crystals, (*a*) NaCl, and (*b*) CsCl, with first-, second-, third-, fourth-, and fifth-neighbor distances

The lattice constant a of the rock-salt crystal is customarily defined as the shortest distance between two similar ions along one of the lattice axes. The nearest neighbors of each Na^+ ion are then six Cl^- ions at a distance $d_1 = a/2$.

[6] E. Madelung, 1918.

The 12 second-neighbor Na^+ ions are located at a distance $d_2 = \sqrt{2}a/2$, the 8 third-neighbor Cl^- ions at $d_3 = \sqrt{3}a/2$, and the 6 fourth-neighbor Na^+ ions at $d_4 = a$. The number of fifth neighbors is 24; these are at a distance $d_5 = \sqrt{5}a/2$. Selecting one Na^+ ion and calculating its contribution to the energy (4.3.19) by performing the summation for its neighbors, we have

$$\varepsilon_{Na^+} = -\frac{\tilde{e}^2}{a}\left[6 - \frac{12}{\sqrt{2}} + \frac{8}{\sqrt{3}} - \frac{6}{\sqrt{4}} + \frac{24}{\sqrt{5}} + \dots\right], \tag{4.3.20}$$

where due account has been taken of the fact that the energy of each ion pair should be shared between the two ions in question.

In its present form the series does not converge at all, as the partial sums oscillate wildly. Evjen proposed a suitable rearrangement of the terms that corresponds to decomposing the crystal into neutral shells. An ion of the rock-salt crystal is singled out; the first layer is then composed of the ions that are on or within the boundaries of a cube of edge a, centered on the selected ion. However, for each ion only the part of its charge that falls within the cube is counted for the layer in question – that is, ions at the face centers, edge centers, and vertices are counted with weights of 1/2, 1/4, and 1/8, respectively. This is shown in Fig. 4.3.

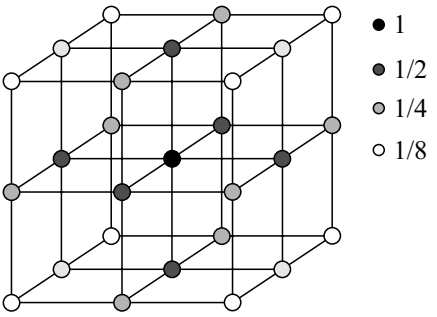


Fig. 4.3. The first neutralizing shell around the central atom, and the weights of the ions within. This shell provides the first approximation in Evjen's method for the calculation of the Madelung energy in a sodium chloride crystal

Using these weights, the contribution of the first shell to the energy is

$$\varepsilon^{(1)}_{Na^+} = -\frac{\tilde{e}^2}{a}\left[\frac{1}{2}6 - \frac{1}{4}\frac{12}{\sqrt{2}} + \frac{1}{8}\frac{8}{\sqrt{3}}\right] = -1.456\,\frac{\tilde{e}^2}{a}\,. \tag{4.3.21}$$

In the next approximation a cube of edge $2a$ is considered around the selected ion. Figure 4.4 shows the ions within an octant of this cube, along with their distances from the central ion. (Once again, the unit is chosen as the distance between nearest neighbors, $d_1 = a/2$.) Ions that lie inside this

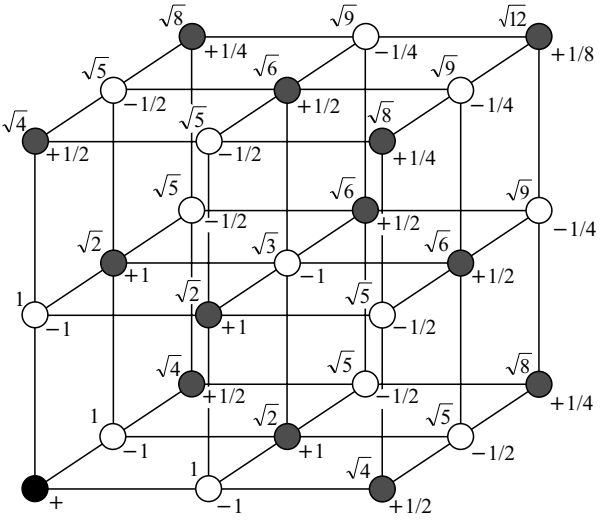


Fig. 4.4. Charges in one octant of the shell around the selected ion (bottom left) that have to be taken into account in the second approximation of Evjen's method. The numbers above and to the left give the distance of these ions from the central one of the shell, while those below and to the right give their weights within this shell

cube are counted with their full charge, while those at the face centers, edge centers, and vertices are counted with weights of 1/2, 1/4, and 1/8, as above.

In this approximation the Madelung energy per sodium ion is given by

$$\begin{aligned}\varepsilon^{(2)}_{\mathrm{Na}^+} &= -\frac{\tilde{e}^2}{a}\left[\left(6-\frac{12}{\sqrt{2}}+\frac{8}{\sqrt{3}}\right)-\frac{1}{2}\left(\frac{6}{\sqrt{4}}-\frac{24}{\sqrt{5}}+\frac{24}{\sqrt{6}}\right)\right.\\ &\qquad\left.+\frac{1}{4}\left(-\frac{12}{\sqrt{8}}+\frac{24}{\sqrt{9}}\right)-\frac{1}{8}\frac{8}{\sqrt{12}}\right]\\ &= -1.752\,\frac{\tilde{e}^2}{a}\,. \end{aligned} \tag{4.3.22}$$

The same applies for each Cl^- ion, too. When expressed in terms of the nearest-neighbor distance d_1 rather than the lattice constant a, the Madelung energy per NaCl molecule is

$$\varepsilon_{\mathrm{NaCl}} = -1.752\,\frac{\tilde{e}^2}{d_1}\,. \tag{4.3.23}$$

Taking the next shell into account would give a correction smaller than one percent. This is the same as the difference with the result obtained from the Ewald summation method, and considered as precise:

$$\varepsilon_{\mathrm{NaCl}} = -1.74756\,\frac{\tilde{e}^2}{d_1}\,. \tag{4.3.24}$$

This value comes from a clever evaluation of the explicit formula

$$\varepsilon_{\mathrm{Na}^+} = \frac{\tilde{e}^2}{a} \sum_{mnp}{}' \frac{(-1)^{m+n+p}}{\sqrt{m^2+n^2+p^2}}\,, \tag{4.3.25}$$

where m, n, and p run through all integers from $-\infty$ to ∞, only the origin $m = n = p = 0$ is excluded.

In the cesium chloride structure Cs^+ and Cl^- ions form interpenetrating cubic lattices, in such a way that the center of each elementary cube formed by Cs^+ is occupied by a Cl^- ion – and vice versa. Thus each Cs^+ ion (cation) is surrounded by 8 Cl^- ions (anions), at a distance $d_1 = \sqrt{3}a/2$. The 6 second-neighbor and 12 third-neighbor cations are found at distances $d_2 = a$ and $d_3 = \sqrt{2}a$. The 16 fourth-neighbor sites, at $d_4 = \sqrt{11}a/2$, are occupied by anions, and the 8 fifth-neighbor sites by cations again, at $d_5 = \sqrt{3}a$. The Madelung energy per Cs^+ ion is now

$$\varepsilon_{\mathrm{Cs}^+} = -\frac{\tilde{e}^2}{2a}\left[\frac{2}{\sqrt{3}}8 - 6 - \frac{12}{\sqrt{2}} + \frac{2}{\sqrt{11}}16 - \frac{8}{\sqrt{3}} + \dots\right]. \tag{4.3.26}$$

Decomposing this series, too, into suitably chosen partial sums,

$$\varepsilon_{\mathrm{Cs}^+} = -2.0354\,\frac{\tilde{e}^2}{2a}\,. \tag{4.3.27}$$

Expressed in terms of the nearest-neighbor distance d_1 rather than the edge a of the cubic lattice, the Madelung energy per CsCl molecule is

$$\varepsilon_{\mathrm{CsCl}} = -1.7627\,\frac{\tilde{e}^2}{d_1}\,. \tag{4.3.28}$$

When the contribution of the Coulomb interactions to the total energy E of a crystal made up of N molecules is expressed with the nearest-neighbor distance as

$$\varepsilon = \frac{E}{N} = -\alpha\,\frac{\tilde{e}^2}{d_1}\,, \tag{4.3.29}$$

the coefficient α is called the *Madelung constant*. This constant is just the ratio of the Coulomb energy per molecule in the crystalline state and the Coulomb energy of a single molecule in which the ions are separated by the nearest-neighbor distance d_1 characteristic of the lattice.

Similarly, the Madelung constant of an ionic crystal $\mathrm{M}^{Z+}\mathrm{X}^{Z-}$ made up of charge-Ze ions is defined in terms of the electrostatic energy of a single $\mathrm{M}^{Z+}\mathrm{X}^{Z-}$ molecule as

$$\frac{E}{N} = -\alpha\,\frac{Z^2\tilde{e}^2}{d_1}\,. \tag{4.3.30}$$

This cannot be unambiguously generalized to the case of ionic crystals made up of cations of charge $Z_1 e$ and anions of charge $-Z_2 e$, i.e., of ionic composition $\mathrm{M}_{n_1}^{Z_1+}\mathrm{X}_{n_2}^{Z_2-}$ – where electric neutrality implies $n_1 Z_1 = n_2 Z_2$. Thus in

standard tables one finds different values for the Madelung constant, corresponding to different definitions. In one commonly used definition the energy per chemical unit is given by

$$\frac{E}{N} = -\alpha \frac{Z_1 Z_2 \tilde{e}^2}{d_1}, \tag{4.3.31}$$

where d_1 is the smallest distance between anions and cations, as before. The corresponding Madelung constants of some typical ionic crystals are listed in Table 4.4. The underlying structures will be presented in Chapter 7.

Table 4.4. The Madelung constant for some typical ionic crystals

Structure			Ionic	Madelung
Name	Notation		composition	constant
Sodium chloride (NaCl)	B1	cF8	M^+X^-	1.74756
Cesium chloride (CsCl)	B2	cP2	M^+X^-	1.76267
Sphalerite (ZnS)	B3	cF8	$M^{2+}X^{2-}$	1.63806
Wurtzite (ZnS)	B4	hP4	$M^{2+}X^{2-}$	1.64132
Fluorite (CaF_2)	C1	cF12	$M^{2+}X_2^-$	2.51939
Cuprite (Cu_2O)	C3	cP6	$M_2^+X^{2-}$	2.22124
Corundum (Al_2O_3)	$D5_1$	hR10	$M_2^{3+}X_3^{2-}$	4.1719

The stability of ionic crystals cannot be understood in terms of the Madelung energy alone. Once again, one has to take into account the repulsion between ion cores that has been discussed for molecular crystals. Instead of the $1/r^{12}$ term in the Lennard-Jones potential, following M. Born and J. E. Mayer (1932), the repulsion is customarily taken into account by an exponentially decreasing term in this case. This is the *Born–Mayer approximation*. Denoting the shortest distance between an anion and a cation by r, the energy per chemical unit can be written as

$$\frac{E}{N} = -\alpha \frac{Z_1 Z_2 \tilde{e}^2}{r} + B \exp(-r/l)\,. \tag{4.3.32}$$

This expression is required to attain its minimum at d_1, the equilibrium distance between neighbors, which can be determined from the equation

$$\alpha \frac{Z_1 Z_2 \tilde{e}^2}{d_1^2} - \frac{B}{l} \exp(-d_1/l) = 0\,. \tag{4.3.33}$$

Using this equation to eliminate B, we have

$$\frac{E}{N} = -\alpha \frac{Z_1 Z_2 \tilde{e}^2}{d_1} \left[1 - \frac{l}{d_1}\right] \tag{4.3.34}$$

for the equilibrium configuration.

The quantity l/d_1 is related to compressibility. Its value is on the order of 0.1. When this correction is also taken into account, good agreement is found between the theoretical values obtained for the cohesive energies of ionic crystals and the experimental data listed in Table 4.2. To calculate the total energy, the electrostatic contribution is not the only one to be taken into account. According to the laws of quantum mechanics, ions are not fixed rigidly to lattice sites but oscillate even at the absolute zero of temperature. However, the energy corrections due to such *zero-point vibrations* do not exceed a few percent.

4.4 Covalent Bond

In solids built up of identical atoms or different atoms with not too disparate electronegativities, the characteristic electron transfer of ionic bonding does not occur. Instead, new electronic states appear, in which electrons are located between two neighboring atoms – and thus belong to both of them. The bond brought about by such electron states is called a *covalent bond*.[7] The electron density reaches its maximum in the region between the two atoms. This is clearly seen in Fig. 4.5, which shows the plane section of electron density contours around two neighboring atoms of a typical covalently bonded solid, a germanium crystal.

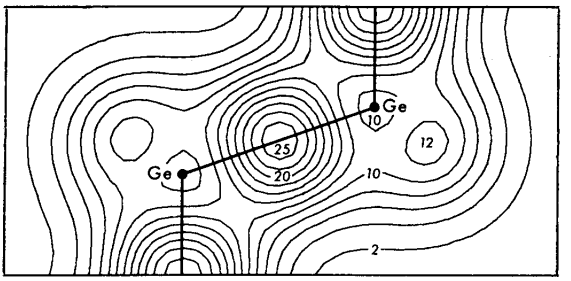


Fig. 4.5. Calculated density contours in the $(1\bar{1}0)$ plane of the electrons forming the covalent bond in germanium. [J. R. Chelikowsky and M. L. Cohen, *Phys. Rev. B* **14**, 556 (1976)]

As a matter of fact, the covalent bond is the generalization of the chemical concept of intramolecular valence bonds to solids. For a mathematical

[7] The terms *homopolar* and *homeopolar bond* are also used in the literature; they refer to situations where the electronegativities of the two atoms are identical and similar, respectively.

description, consider first a diatomic molecule that has two electrons (H_2). The position of atoms A and B will be denoted by $\boldsymbol{R}_\mathrm{A}$ and $\boldsymbol{R}_\mathrm{B}$, and we shall assume that each atom contains one of the two electrons. This assumption is justified when the atoms are sufficiently separated, and so the electron clouds do not overlap. We shall now ask: What happens when the two atoms approach one another so much that the other's presence is no longer a weak perturbation?

Using the notations $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ for the position vectors of the two electrons, the Hamiltonian of the system reads

$$\mathcal{H} = -\frac{\hbar^2}{2m_\mathrm{e}}\boldsymbol{\nabla}_1^2 - \frac{\hbar^2}{2m_\mathrm{e}}\boldsymbol{\nabla}_2^2 - \sum_{i=1}^{2}\frac{\tilde{e}^2}{|\boldsymbol{r}_i - \boldsymbol{R}_\mathrm{A}|} - \sum_{i=1}^{2}\frac{\tilde{e}^2}{|\boldsymbol{r}_i - \boldsymbol{R}_\mathrm{B}|} + \frac{\tilde{e}^2}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|} + \frac{\tilde{e}^2}{|\boldsymbol{R}_\mathrm{A} - \boldsymbol{R}_\mathrm{B}|}\,. \tag{4.4.1}$$

This Hamiltonian does not usually admit an analytical solution to the Schrödinger equation. The customary way of tackling this problem consists of choosing a trial wavefunction (ansatz) with a few variational parameters which are determined from the condition that energy should take a minimum value. The individual methods differ in the assumptions made about the form of the multi-electron wavefunction.

4.4.1 The Valence-Bond Method

Even simpler than the variational method is the *Heitler–London approximation*,[8] also known as the *valence-bond method*. This approximation is based on atomic functions. The underlying assumption is that one of the electrons forming the covalent bond is mostly found in the vicinity of one nucleus, while the other electron around the other nucleus. Then, following the same steps as for the van der Waals interaction, the Hamiltonian is split into three parts. As the whole system is comprised of only two atoms, we have

$$\mathcal{H}(\boldsymbol{r}_1, \boldsymbol{r}_2) = \mathcal{H}_\mathrm{A}(\boldsymbol{r}_1) + \mathcal{H}_\mathrm{B}(\boldsymbol{r}_2) + \mathcal{H}_\mathrm{int}(\boldsymbol{r}_1, \boldsymbol{r}_2)\,, \tag{4.4.2}$$

where

$$\begin{aligned}\mathcal{H}_\mathrm{A}(\boldsymbol{r}_1) &= -\frac{\hbar^2}{2m_\mathrm{e}}\boldsymbol{\nabla}_1^2 - \frac{\tilde{e}^2}{|\boldsymbol{r}_1 - \boldsymbol{R}_\mathrm{A}|}\,,\\ \mathcal{H}_\mathrm{B}(\boldsymbol{r}_2) &= -\frac{\hbar^2}{2m_\mathrm{e}}\boldsymbol{\nabla}_2^2 - \frac{\tilde{e}^2}{|\boldsymbol{r}_2 - \boldsymbol{R}_\mathrm{B}|}\,,\end{aligned} \tag{4.4.3}$$

and

$$\mathcal{H}_\mathrm{int}(\boldsymbol{r}_1, \boldsymbol{r}_2) = -\frac{\tilde{e}^2}{|\boldsymbol{r}_1 - \boldsymbol{R}_\mathrm{B}|} - \frac{\tilde{e}^2}{|\boldsymbol{r}_2 - \boldsymbol{R}_\mathrm{A}|} + \frac{\tilde{e}^2}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|} + \frac{\tilde{e}^2}{|\boldsymbol{R}_\mathrm{A} - \boldsymbol{R}_\mathrm{B}|}\,. \tag{4.4.4}$$

[8] W. Heitler and F. London (1927).

Assuming that the solutions to the one-center Schrödinger equation

$$\mathcal{H}_{\mathrm{A}}(\boldsymbol{r}_1)\psi_{\mathrm{A}}(\boldsymbol{r}_1) = \varepsilon_{\mathrm{A}}\psi_{\mathrm{A}}(\boldsymbol{r}_1) \tag{4.4.5}$$

and to its counterpart for the wavefunction $\psi_{\mathrm{B}}(\boldsymbol{r}_2)$ of the other electron around atom B are known – e.g., for a 1s state it is

$$\psi_{\mathrm{A}}(\boldsymbol{r}) = \left(\frac{1}{\pi a_0^3}\right)^{1/2} \mathrm{e}^{-|\boldsymbol{r}-\boldsymbol{R}_{\mathrm{A}}|/a_0}\,, \tag{4.4.6}$$

the wavefunction of the two-electron system can be written as the product of one-particle wavefunctions,

$$\psi(\boldsymbol{r}_1, \boldsymbol{r}_2) = \psi_{\mathrm{A}}(\boldsymbol{r}_1)\psi_{\mathrm{B}}(\boldsymbol{r}_2)\,. \tag{4.4.7}$$

This automatically satisfies the two-center Schrödinger equation

$$\Big[\mathcal{H}_{\mathrm{A}}(\boldsymbol{r}_1) + \mathcal{H}_{\mathrm{B}}(\boldsymbol{r}_2)\Big]\psi(\boldsymbol{r}_1, \boldsymbol{r}_2) = \big(\varepsilon_{\mathrm{A}} + \varepsilon_{\mathrm{B}}\big)\psi(\boldsymbol{r}_1, \boldsymbol{r}_2)\,. \tag{4.4.8}$$

When the expectation value of the interaction Hamiltonian is calculated with this wavefunction, an estimate is obtained for the binding energy.

Before doing so, it should be noted that the wavefunction (4.4.7) does not satisfy the general condition of antisymmetry required for fermionic systems. This lack of antisymmetry is the consequence of the fact that the two electrons were considered as distinguishable, thus $\mathcal{H}_{\mathrm{A}}(\boldsymbol{r}_1)+\mathcal{H}_{\mathrm{B}}(\boldsymbol{r}_2)$ is not invariant under the interchange of their coordinates – even though the total Hamiltonian is. Therefore another wavefunction has to be chosen, one that is consistent with the Pauli exclusion principle. To see antisymmetry clearly, electron spins have to be taken into account, too, even though the interaction is spin independent. Using the notations $s = \pm$, or $\uparrow$ and $\downarrow$ (rather than $m_s = \pm\frac{1}{2}$) for the two possible values of the spin variable s of the electron, the corresponding wavefunctions are

$$\chi(\uparrow) = |\uparrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \text{and} \qquad \chi(\downarrow) = |\downarrow\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{4.4.9}$$

In the nonrelativistic case when the spin–orbit interaction is neglected, the spin independence of the Hamiltonian permits the separation of spatial and spin variables in the wavefunction:

$$\Psi(\boldsymbol{r}_1, s_1, \boldsymbol{r}_2, s_2) = \psi(\boldsymbol{r}_1, \boldsymbol{r}_2)\chi(s_1, s_2)\,. \tag{4.4.10}$$

Either the wavefunction is symmetric in its spatial variables:

$$\psi_{\mathrm{s}}(\boldsymbol{r}_1, \boldsymbol{r}_2) = N_+\big[\psi_{\mathrm{A}}(\boldsymbol{r}_1)\psi_{\mathrm{B}}(\boldsymbol{r}_2) + \psi_{\mathrm{A}}(\boldsymbol{r}_2)\psi_{\mathrm{B}}(\boldsymbol{r}_1)\big]\,, \tag{4.4.11}$$

and then it has to be antisymmetric – singlet – in its spin variables:

$$\chi_{\mathrm{s}}(s_1, s_2) = \tfrac{1}{\sqrt{2}}\big[|\uparrow\rangle_1 |\downarrow\rangle_2 - |\downarrow\rangle_1 |\uparrow\rangle_2\big]\,, \tag{4.4.12}$$

or it is antisymmetric in its spatial variables:

$$\psi_{\mathrm{t}}(\boldsymbol{r}_1, \boldsymbol{r}_2) = N_-\big[\psi_{\mathrm{A}}(\boldsymbol{r}_1)\psi_{\mathrm{B}}(\boldsymbol{r}_2) - \psi_{\mathrm{A}}(\boldsymbol{r}_2)\psi_{\mathrm{B}}(\boldsymbol{r}_1)\big]\,, \tag{4.4.13}$$

and then it has to be symmetric – triplet – in its spin variables:

$$\chi_{\mathrm{t}}(s_1, s_2) = \begin{cases} |\uparrow\rangle_1 |\uparrow\rangle_2 \\ \tfrac{1}{\sqrt{2}}\big[|\uparrow\rangle_1 |\downarrow\rangle_2 + |\downarrow\rangle_1 |\uparrow\rangle_2\big] \\ |\downarrow\rangle_1 |\downarrow\rangle_2\,. \end{cases} \tag{4.4.14}$$

In what follows, we shall suppress the spin-dependent parts, as the Hamiltonian does not act on the spins. We shall nevertheless bear in mind that the spin state of ψ_{s} is a singlet, while that of ψ_{t} is a triplet. We now have to determine the constant $N_\pm$ that takes care of the proper normalization of the wavefunctions. Since the wavefunctions around different ion cores are not mutually orthogonal, the overlap integral

$$S_{\mathrm{AB}} = \int \psi_{\mathrm{A}}^*(\boldsymbol{r})\psi_{\mathrm{B}}(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r} \tag{4.4.15}$$

must be taken into account. The normalization factor is then

$$N_\pm = \frac{1}{\sqrt{2(1 \pm |S_{\mathrm{AB}}|^2)}}\,. \tag{4.4.16}$$

In the Heitler–London approximation these completely antisymmetrized wavefunctions are chosen, and the energy is then calculated as the expectation value of the total Hamiltonian,

$$\begin{aligned} \varepsilon_{\mathrm{s}} &= \langle\psi_{\mathrm{s}}|\mathcal{H}|\psi_{\mathrm{s}}\rangle = \varepsilon_{\mathrm{A}} + \varepsilon_{\mathrm{B}} + 2N_+^2(C + I)\,, \\ \varepsilon_{\mathrm{t}} &= \langle\psi_{\mathrm{t}}|\mathcal{H}|\psi_{\mathrm{t}}\rangle = \varepsilon_{\mathrm{A}} + \varepsilon_{\mathrm{B}} + 2N_-^2(C - I)\,, \end{aligned} \tag{4.4.17}$$

where

$$\begin{aligned} C &= \int \psi_{\mathrm{A}}^*(\boldsymbol{r}_1)\psi_{\mathrm{B}}^*(\boldsymbol{r}_2)\mathcal{H}_{\mathrm{int}}\psi_{\mathrm{A}}(\boldsymbol{r}_1)\psi_{\mathrm{B}}(\boldsymbol{r}_2)\,\mathrm{d}\boldsymbol{r}_1\,\mathrm{d}\boldsymbol{r}_2\,, \\ I &= \int \psi_{\mathrm{A}}^*(\boldsymbol{r}_1)\psi_{\mathrm{B}}^*(\boldsymbol{r}_2)\mathcal{H}_{\mathrm{int}}\psi_{\mathrm{B}}(\boldsymbol{r}_1)\psi_{\mathrm{A}}(\boldsymbol{r}_2)\,\mathrm{d}\boldsymbol{r}_1\,\mathrm{d}\boldsymbol{r}_2\,. \end{aligned} \tag{4.4.18}$$

C is the direct contribution of the interaction (Coulomb integral), while I, the *exchange integral* comes from quantum mechanical exchange, the existence of which was pointed out independently by Heisenberg[9] and Dirac[10] in 1926.

[9] Werner Karl Heisenberg (1901–1976) was awarded the Nobel Prize in 1932 "for the creation of quantum mechanics, the application of which has, inter alia, led to the discovery of the allotropic forms of hydrogen".

[10] See footnote on page 32.

If there were no overlap between the wavefunctions of electrons on neighboring ion cores then the two electrons could not be exchanged and I would vanish. When the overlap is finite and the wavefunctions are orthogonal (this would be the case for the two electrons on different orbitals around the same atom) then I can be shown to be positive. This implies Hund's first rule: within an atom the orthogonal orbitals of the same azimuthal quantum number are occupied by the electrons in such a way that the total spin should be maximal. For electrons on neighboring atoms the lack of orthogonality means that the exchange integral I can be negative. The sequence of the levels is determined by the relation of I and CS_{AB}^2, as

$$\varepsilon_{\mathrm{s}} > \varepsilon_{\mathrm{t}}\,, \qquad \text{if} \qquad \frac{C+I}{1+S_{\mathrm{AB}}^2} > \frac{C-I}{1-S_{\mathrm{AB}}^2}\,, \tag{4.4.19}$$

whence

$$I > CS_{\mathrm{AB}}^2\,. \tag{4.4.20}$$

Usually $S_{\mathrm{AB}} \ll 1$, in which case the sign of I will determine whether the triplet or the singlet state is of lower energy.

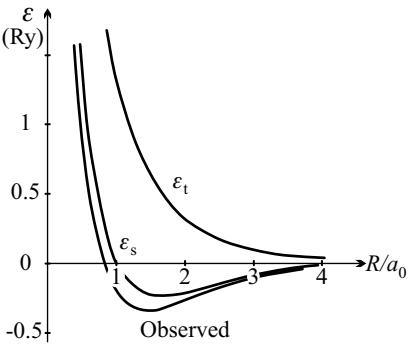


Fig. 4.6. Binding energies of singlet and triplet states in the Heitler–London approximation as functions of the separation R of the nuclei

Heitler and London's calculations have shown that I is strongly negative for the hydrogen molecule. Figure 4.6 shows the energies of the singlet and triplet states as functions of the separation $R = |\boldsymbol{R}_{\mathrm{A}} - \boldsymbol{R}_{\mathrm{B}}|$ of the nuclei. For large values of R the energy of free atoms is recovered. The short-distance behavior is, on the other hand, dominated by repulsion that is rooted in the exclusion principle. In the singlet state the energy has its minimum at an intermediate value of the separation R, giving rise to bonding. In this approach a binding energy of 3.14 eV and a bonding distance of 0.87 Å are found. Despite the roughness of the approximation, these values are surprisingly close to the experimental data, 4.75 eV and 0.74 Å. For our purposes, however, the most important point in this result is not this fair agreement (as it is not

difficult to find better approximations for the binding energy) but that it highlights the energy difference between singlet and triplet states in a case when the Hamiltonian is spin independent. The energy difference is entirely due to the symmetric and antisymmetric characters of the spatial parts of the wavefunctions, and the exchange integral I may lead to the appearance of magnetic states. We shall return to this point in Chapter 14 where we shall study magnetically ordered systems.

4.4.2 Polar Covalent Bond

The Heitler–London approximation deals only with so-called *valence-bond configurations* – in which one electron is around core A and the other is around core B. The spatial part of the corresponding singlet covalent wavefunction is

$$\psi_{\mathrm{cov}}(\boldsymbol{r}_1, \boldsymbol{r}_2) = \psi_{\mathrm{A}}(\boldsymbol{r}_1)\psi_{\mathrm{B}}(\boldsymbol{r}_2) + \psi_{\mathrm{B}}(\boldsymbol{r}_1)\psi_{\mathrm{A}}(\boldsymbol{r}_2)\,. \tag{4.4.21}$$

During the motion of the electrons it may happen that both electrons are on orbitals around the same atomic core. Such configurations are called *ionic configurations*. When the two nuclei are identical, and therefore electrons are equally likely to be around either atom, the spatial part of the corresponding wavefunction reads

$$\psi_{\mathrm{ion}}(\boldsymbol{r}_1, \boldsymbol{r}_2) = \psi_{\mathrm{A}}(\boldsymbol{r}_1)\psi_{\mathrm{A}}(\boldsymbol{r}_2) \pm \psi_{\mathrm{B}}(\boldsymbol{r}_1)\psi_{\mathrm{B}}(\boldsymbol{r}_2)\,. \tag{4.4.22}$$

The symmetric character of the spatial part implies that the spin variables are, again, in a singlet configuration. To obtain a more precise description than that provided by the Heitler–London approximation for the problem of two electrons moving in the field of two nuclei, one has to take into account not only the covalent terms but also the above ionic one – with a weight that will be determined later. This is known as the *polar covalent bond* or the *ionic–covalent bond*. Then the spatial part of the singlet wavefunction of the two-electron system is chosen as

$$\begin{aligned}\psi(\boldsymbol{r}_1, \boldsymbol{r}_2) = N\big[&\psi_{\mathrm{A}}(\boldsymbol{r}_1)\psi_{\mathrm{B}}(\boldsymbol{r}_2) + \psi_{\mathrm{B}}(\boldsymbol{r}_1)\psi_{\mathrm{A}}(\boldsymbol{r}_2) \\ &+ \lambda\psi_{\mathrm{A}}(\boldsymbol{r}_1)\psi_{\mathrm{A}}(\boldsymbol{r}_2) + \lambda\psi_{\mathrm{B}}(\boldsymbol{r}_1)\psi_{\mathrm{B}}(\boldsymbol{r}_2)\big]\,,\end{aligned} \tag{4.4.23}$$

where λ is a variational parameter. Its value is determined by the minimum energy condition. Calculations have shown that in a hydrogen molecule the energy minimum at equilibrium distance occurs for $\lambda = 0.25$ – that is, the bond is ionic only to the extent of a few percent: it is predominantly covalent. The binding energy is found to be 3.23 eV.

Note that for even more precise results one has to include variational parameters in the atomic wavefunction as well. For example, in the case of atomic s states, the expression

$$\psi(r) = \left(\frac{\gamma^3}{\pi a_0^3}\right)^{1/2} \mathrm{e}^{-\gamma r/a_0} \tag{4.4.24}$$

can be used instead of (4.4.6), and energy is minimized with respect to γ. Calculations show that the minimum occurs for $\gamma > 1$ – that is, because of bonding, electrons approach the nucleus more than in a free atom.

For identical atoms this bonding gives rise to apolar molecules. In solids made up of atoms with different electronegativity values the bond can by no means be purely covalent. Electrons are more attracted to one kind of atom. This is illustrated in Fig. 4.7, where the plane section of the electron density contours are shown in two crystals, GaP and ZnSe. These compounds are made up of elements in groups 13 (IIIA, boron group) and 15 (VA, nitrogen group), and groups 12 (IIB) and 16 (VIA) of the periodic table, respectively. A comparison with Fig. 4.5 clearly shows a more and more pronounced asymmetry of the electron density.

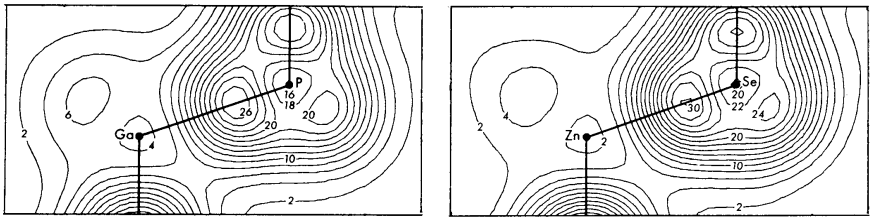


Fig. 4.7. Calculated density contours between two neighboring ions in a plane section of GaP and ZnSe crystals. [J. R. Chelikowsky and M. L. Cohen, *Phys. Rev. B* **14**, 556 (1976)]

When the two atoms are not identical – and thus the symmetry with respect to the exchange of atoms A and B is broken –, the configurations in which both electrons are found around core A and core B appear in the wavefunction with different weight factors. The wavefunction that corresponds to the singlet state reads

$$\begin{aligned}\psi(\boldsymbol{r}_1,\boldsymbol{r}_2) = N\big[&\psi_{\mathrm{A}}(\boldsymbol{r}_1)\psi_{\mathrm{B}}(\boldsymbol{r}_2) + \psi_{\mathrm{B}}(\boldsymbol{r}_1)\psi_{\mathrm{A}}(\boldsymbol{r}_2) \\ &+ \lambda_{\mathrm{A}}\psi_{\mathrm{A}}(\boldsymbol{r}_1)\psi_{\mathrm{A}}(\boldsymbol{r}_2) + \lambda_{\mathrm{B}}\psi_{\mathrm{B}}(\boldsymbol{r}_1)\psi_{\mathrm{B}}(\boldsymbol{r}_2)\big]\,.\end{aligned} \tag{4.4.25}$$

Treating the quantities λ_{A} and λ_{B} as variational parameters, the obtained energy will be lower than that calculated from the Heitler–London wavefunction ψ_{cov} in (4.4.21) – which corresponds to a purely covalent bond. The binding energy is written as

$$E = E_{\mathrm{cov}} - \Delta\,, \tag{4.4.26}$$

where E_{cov} would be the energy of a purely covalent bond, and Δ is called the contribution of ionic resonance. The energy of the covalent bonds A–B between atoms A and B can be estimated by the arithmetic (or in some references, geometric) mean of the bond energies A–A and B–B. The contribution of ionic resonance is a function of the difference of the electronegativities. Thus a simple estimate can be given for the bond energy.

4.4.3 The Molecular-Orbital Method

In the previously presented methods the Schrödinger equation is first solved in the field of one nucleus, and then the obtained one-center electron states are used to construct the wavefunction of the electrons in the multi-atom system. A different approach was proposed by F. Hund and R. S. Mulliken[11] in 1928. First the one-electron states of the multi-atom molecule are determined, and then these are populated with electrons step by step. For this reason the Hund–Mulliken approach is called the *molecular-orbital method (MO method)* nowadays. When calculating the one-electron states some care must be exercised as the potential felt by the electrons is changed by the addition of a new electron – i.e., the potential itself needs to be determined in a self-consistent manner. The simplest method for this is the Hartree–Fock approximation, which will be discussed in Chapter 28. As in the present chapter we just want to develop a qualitative picture about the nature of different types of bonds, we shall not analyze the condition of self-consistency. Instead, we shall assume that for any selected electron the effect of the other electron(s) can be accounted for by using an average potential.

Coulomb repulsion between two immobile cores of a two-atom molecule must be taken into account when calculating the total energy but it can be ignored when the energy of the system of electrons is considered. In this approximation the Hamiltonian is split into two independent parts,

$$\mathcal{H} = \mathcal{H}_1 + \mathcal{H}_2 \,, \tag{4.4.27}$$

where

$$\mathcal{H}_l = -\frac{\hbar^2}{2m_\mathrm{e}}\boldsymbol{\nabla}_l^2 - \frac{\tilde{e}^2}{|\boldsymbol{r}_l - \boldsymbol{R}_\mathrm{A}|} - \frac{\tilde{e}^2}{|\boldsymbol{r}_l - \boldsymbol{R}_\mathrm{B}|} + U_\mathrm{eff}(\boldsymbol{r}_l) \,, \quad l = 1, 2 \,. \tag{4.4.28}$$

The average potential due to the other electron(s) is included in the term $U_\mathrm{eff}(\boldsymbol{r}_l)$.

The state associated with the eigenfunction $\psi_i(\boldsymbol{r})$ of the two-center, one-electron Schrödinger equation

$$\left[-\frac{\hbar^2}{2m_\mathrm{e}}\boldsymbol{\nabla}^2 - \frac{\tilde{e}^2}{|\boldsymbol{r} - \boldsymbol{R}_\mathrm{A}|} - \frac{\tilde{e}^2}{|\boldsymbol{r} - \boldsymbol{R}_\mathrm{B}|} + U_\mathrm{eff}(\boldsymbol{r})\right] \psi_i(\boldsymbol{r}) = \varepsilon_i \psi_i(\boldsymbol{r}) \tag{4.4.29}$$

is called the ith *molecular orbital*. Solving this two-center Schrödinger equation usually requires further approximations. We shall present one possibility in the next subsection.

Before doing so it should be mentioned that, provided the molecular orbitals are known, the separation of the Hamiltonian permits us to write the

[11] Robert Sanderson Mulliken (1896–1986) was awarded the Nobel Prize in Chemistry in 1966 "for his fundamental work concerning chemical bonds and the electronic structure of molecules by the molecular orbital method".

total wavefunction of the two-electron system as the product of the wavefunctions of single electrons on the molecular orbitals,

$$\psi(\boldsymbol{r}_1, \boldsymbol{r}_2) = \psi_i(\boldsymbol{r}_1)\psi_j(\boldsymbol{r}_2)\,. \tag{4.4.30}$$

This wavefunction must be complemented by a spin-dependent part, and consistency with the Pauli exclusion principle must be ensured. If the same molecular orbital ψ_i is selected for both electrons, the spin part must be chosen antisymmetric. When different molecular orbitals, and thus different spatial functions are selected, there are two options. Either the symmetric combination of the variables $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$,

$$\psi(\boldsymbol{r}_1, \boldsymbol{r}_2) = \frac{1}{\sqrt{2}}\big[\psi_i(\boldsymbol{r}_1)\psi_j(\boldsymbol{r}_2) + \psi_i(\boldsymbol{r}_2)\psi_j(\boldsymbol{r}_1)\big] \tag{4.4.31}$$

is chosen, and then the spin part is given by an antisymmetric (singlet) wavefunction, or the antisymmetric combination

$$\psi(\boldsymbol{r}_1, \boldsymbol{r}_2) = \frac{1}{\sqrt{2}}\big[\psi_i(\boldsymbol{r}_1)\psi_j(\boldsymbol{r}_2) - \psi_i(\boldsymbol{r}_2)\psi_j(\boldsymbol{r}_1)\big]\,, \tag{4.4.32}$$

which is then complemented by a symmetric spin part.

4.4.4 The LCAO Method

An approximate form of the molecular orbitals can be obtained using atomic wavefunctions. Assuming that the solutions $\psi(\boldsymbol{r})$ to the atomic Schrödinger equation (4.4.5) are known, following the proposal of J. E. Lennard-Jones (1929), the molecular orbital is constructed as the linear combination of the solutions $\psi_{\mathrm{A}}(\boldsymbol{r})$ and $\psi_{\mathrm{B}}(\boldsymbol{r})$ obtained for individual atoms,

$$\psi_i(\boldsymbol{r}) = c_{\mathrm{A}}\psi_{\mathrm{A}}(\boldsymbol{r}) + c_{\mathrm{B}}\psi_{\mathrm{B}}(\boldsymbol{r})\,. \tag{4.4.33}$$

This is the *linear combination of atomic orbitals* (LCAO) method.

If the atoms are identical, the two atomic wavefunctions are expected to appear with the same probability, i.e., $c_{\mathrm{A}} = \pm c_{\mathrm{B}}$. This is indeed the case: the combination is either symmetric or antisymmetric. When the overlap integral defined in (4.4.15) is taken into consideration, the normalized wavefunction is

$$\psi_{\pm}(\boldsymbol{r}) = \frac{1}{\sqrt{2 \pm (S_{\mathrm{AB}} + S^*_{\mathrm{AB}})}}\left[\psi_{\mathrm{A}}(\boldsymbol{r}) \pm \psi_{\mathrm{B}}(\boldsymbol{r})\right], \tag{4.4.34}$$

and the corresponding energy value is

$$\varepsilon_{\pm} = \frac{\varepsilon_{\mathrm{A}} + \varepsilon_{\mathrm{B}} \pm (\varepsilon_{\mathrm{AB}} + \varepsilon^*_{\mathrm{AB}})}{2 \pm (S_{\mathrm{AB}} + S^*_{\mathrm{AB}})}\,, \tag{4.4.35}$$

where

$$\varepsilon_{\mathrm{A}} = \int \psi_{\mathrm{A}}^{*}(\boldsymbol{r}) \mathcal{H}_1 \psi_{\mathrm{A}}(\boldsymbol{r})\, \mathrm{d}\boldsymbol{r}\,,$$
$$\varepsilon_{\mathrm{B}} = \int \psi_{\mathrm{B}}^{*}(\boldsymbol{r}) \mathcal{H}_1 \psi_{\mathrm{B}}(\boldsymbol{r})\, \mathrm{d}\boldsymbol{r}\,, \tag{4.4.36}$$
$$\varepsilon_{\mathrm{AB}} = \int \psi_{\mathrm{A}}^{*}(\boldsymbol{r}) \mathcal{H}_1 \psi_{\mathrm{B}}(\boldsymbol{r})\, \mathrm{d}\boldsymbol{r}\,.$$

The molecular energy levels are shifted with respect to the atomic levels ε_{A} and ε_{B}.

If the overlap is ignored, the amount of splitting is determined by the resonance integral $\varepsilon_{\mathrm{AB}}$. This is generally negative since it contains the matrix element of the Coulomb attraction between the nucleus and the electron. Then the spatially symmetric combination has the lowest energy. This configuration is called a *bonding state*, while the spatially antisymmetric combination gives rise to an *antibonding state*. The wavefunctions of the bonding and antibonding molecular orbitals and the corresponding electron densities are shown in Fig. 4.8 when a hydrogenic 1s wavefunction is chosen as the atomic function.

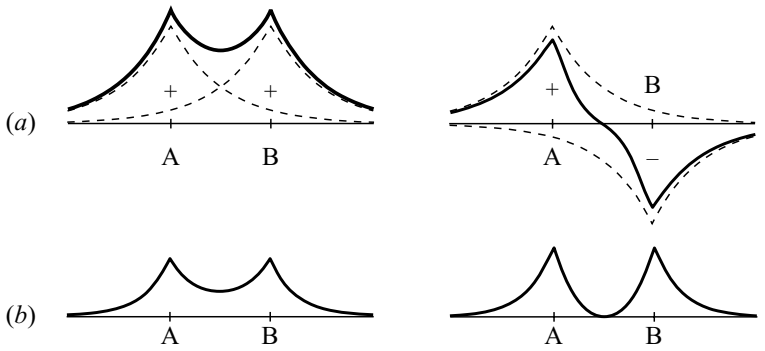


Fig. 4.8. (*a*) The wavefunction of the electrons for bonding (left) and antibonding (right) molecular orbitals in a two-atom system. (*b*) The corresponding charge distributions

In bonding states electron density is large between the two atoms: bonding localizes the electrons to this region. (Figure 4.5 shows this for germanium.) In antibonding orbitals the opposite happens: electron density is smaller in this region than it would be for free atoms at A and B. Therefore one can say that covalent bonding is caused by electron pairs that are located between the atoms with a high probability.

The bonding state formed by the s states is cylindrically symmetric around the axis joining the two atoms, and also symmetric with respect to inversion in the midpoint. It is also called the σ_{g} state. The antibonding state is antisymmetric with respect to space inversion, hence it is also known as the σ_{u}

state. Atomic p orbitals give rise to π molecular orbitals. These can also be classified according to their symmetry properties.

Since the wavefunctions $\psi_{\mathrm{A}}(\boldsymbol{r})$ and $\psi_{\mathrm{B}}(\boldsymbol{r})$ that correspond to electron states around ion cores A and B contain in their arguments the variables $\boldsymbol{r} - \boldsymbol{R}_{\mathrm{A}}$ and $\boldsymbol{r} - \boldsymbol{R}_{\mathrm{B}}$, the quantities $\varepsilon_{\mathrm{AB}}$ and S_{AB} as well as the energies of the bonding and antibonding states depend on the distance between the two atoms. This is shown in Fig. 4.9. The equilibrium distance between the two atoms is determined by the location of the energy minimum of the bonding state.

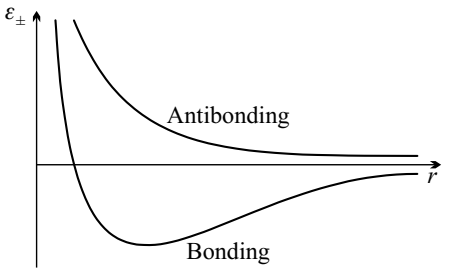


Fig. 4.9. Variation of the energy of bonding and antibonding electron states with the separation of the two atoms

In the ground state of a hydrogen molecule the two electrons occupy the bonding state with opposite spins, making a spin singlet; antibonding states remain empty. This way a saturated covalent bond is formed. When the number of electrons in bonding states is less than the maximum permitted by the Pauli exclusion principle, the bond is unsaturated. On the other hand, when the number of electrons exceeds the number of bonding states, some antibonding states need to be occupied. Then the energy of the resulting state may be so high that no bond is formed. That is why the He_2 molecule is not stable.

When the molecular-orbital method is used in conjunction with the LCAO approximation for a hydrogen molecule, the obtained results are worse than those of the Heitler–London approximation. Although the calculated equilibrium distance of the nuclei, 0.85 Å, is closer to the observed value, the computed binding energy, 2.68 eV is far too low. The reason for this is easy to understand. When both electrons are accommodated on the symmetric binding orbital with opposite spins, the spatial part of the two-particle wavefunction may also be written as

$$\begin{aligned}\psi(\boldsymbol{r}_1, \boldsymbol{r}_2) = \frac{1}{2 + S_{\mathrm{AB}} + S_{\mathrm{AB}}^*} \big[&\psi_{\mathrm{A}}(\boldsymbol{r}_1)\psi_{\mathrm{B}}(\boldsymbol{r}_2) + \psi_{\mathrm{B}}(\boldsymbol{r}_1)\psi_{\mathrm{A}}(\boldsymbol{r}_2) \\ &+ \psi_{\mathrm{A}}(\boldsymbol{r}_1)\psi_{\mathrm{A}}(\boldsymbol{r}_2) + \psi_{\mathrm{B}}(\boldsymbol{r}_1)\psi_{\mathrm{B}}(\boldsymbol{r}_2)\big] .\end{aligned} \tag{4.4.37}$$

It is readily seen that covalent and ionic configurations have the same weight in this approximation. The exceedingly large weight of the ionic

contribution leads to a worse result than that obtained in the Heitler–London approximation, where only covalent configurations are taken into account. As we have seen, better results can be obtained when the relative weight of the two types of configuration is also a free parameter.

4.4.5 Molecular Orbitals Between Different Atoms

While the bond in a hydrogen molecule (H_2) is dominantly covalent, this is not at all the case for a hydrogen fluoride molecule (HF). Here the highly electronegative fluorine atom pulls the electron from the hydrogen – nonetheless the bond will not be perfectly ionic, only to about 50%. Using the molecular-orbital method in the LCAO approximation, the wavefunction of the molecular orbital is again chosen as

$$\psi(\boldsymbol{r}) = c_{\mathrm{A}}\psi_{\mathrm{A}}(\boldsymbol{r}) + c_{\mathrm{B}}\psi_{\mathrm{B}}(\boldsymbol{r}) \tag{4.4.38}$$

but c_{A} and c_{B} will be different. The Schrödinger equation $\mathcal{H}\psi = \varepsilon\psi$ can be written in the equivalent form

$$\sum_{\nu} c_{\nu}\left(H_{\mu\nu} - S_{\mu\nu}\varepsilon\right) = 0 \qquad \mu,\nu = \mathrm{A},\,\mathrm{B}\,, \tag{4.4.39}$$

where

$$\begin{aligned} H_{\mu\nu} &= \int \psi_{\mu}^{*}(\boldsymbol{r})\mathcal{H}\psi_{\nu}(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r}\,, \\ S_{\mu\nu} &= \int \psi_{\mu}^{*}(\boldsymbol{r})\psi_{\nu}(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r}\,. \end{aligned} \tag{4.4.40}$$

The above equation has nontrivial solutions only when

$$|H_{\mu\nu} - S_{\mu\nu}\varepsilon| = 0\,. \tag{4.4.41}$$

The energy of the molecular orbital can be determined from this equation. The lower (higher) energy solution corresponds to the bonding (antibonding) state.

Note that when both electrons are in bonding states, the wavefunction

$$\psi(\boldsymbol{r}_1,\boldsymbol{r}_2) = [c_{\mathrm{A}}\psi_{\mathrm{A}}(\boldsymbol{r}_1) + c_{\mathrm{B}}\psi_{\mathrm{B}}(\boldsymbol{r}_1)][c_{\mathrm{A}}\psi_{\mathrm{A}}(\boldsymbol{r}_2) + c_{\mathrm{B}}\psi_{\mathrm{B}}(\boldsymbol{r}_2)] \tag{4.4.42}$$

is very similar to (4.4.25). However, the two expressions are usually different, since (4.4.25) has two fitting parameters, λ_{A} and λ_{B}, while only one is left in (4.4.42) after normalization. The MO method can yield the same results only if electrons are allowed to occupy the orthogonal antibonding orbital with a weight λ. This way of taking the antibonding states into account is called *configuration interaction* (CI) in quantum chemistry. Quantitatively correct results can be obtained only by including this interaction.

4.4.6 Slater Determinant Form of the Wavefunction

For future generalization it will be useful to write the two-particle states built up of atomic states as a Slater determinant.[12] To start with, one constructs four functions using the atomic wavefunctions ψ_{A} and ψ_{B}:

$$\Psi_1 = \frac{1}{\sqrt{2}} \begin{vmatrix} \psi_{\mathrm{A}}(\boldsymbol{r}_1)|\uparrow\rangle_1 & \psi_{\mathrm{B}}(\boldsymbol{r}_1)|\uparrow\rangle_1 \\ \psi_{\mathrm{A}}(\boldsymbol{r}_2)|\uparrow\rangle_2 & \psi_{\mathrm{B}}(\boldsymbol{r}_2)|\uparrow\rangle_2 \end{vmatrix}, \tag{4.4.43-a}$$

$$\Psi_2 = \frac{1}{\sqrt{2}} \begin{vmatrix} \psi_{\mathrm{A}}(\boldsymbol{r}_1)|\uparrow\rangle_1 & \psi_{\mathrm{B}}(\boldsymbol{r}_1)|\downarrow\rangle_1 \\ \psi_{\mathrm{A}}(\boldsymbol{r}_2)|\uparrow\rangle_2 & \psi_{\mathrm{B}}(\boldsymbol{r}_2)|\downarrow\rangle_2 \end{vmatrix}, \tag{4.4.43-b}$$

$$\Psi_3 = \frac{1}{\sqrt{2}} \begin{vmatrix} \psi_{\mathrm{A}}(\boldsymbol{r}_1)|\downarrow\rangle_1 & \psi_{\mathrm{B}}(\boldsymbol{r}_1)|\uparrow\rangle_1 \\ \psi_{\mathrm{A}}(\boldsymbol{r}_2)|\downarrow\rangle_2 & \psi_{\mathrm{B}}(\boldsymbol{r}_2)|\uparrow\rangle_2 \end{vmatrix}, \tag{4.4.43-c}$$

$$\Psi_4 = \frac{1}{\sqrt{2}} \begin{vmatrix} \psi_{\mathrm{A}}(\boldsymbol{r}_1)|\downarrow\rangle_1 & \psi_{\mathrm{B}}(\boldsymbol{r}_1)|\downarrow\rangle_1 \\ \psi_{\mathrm{A}}(\boldsymbol{r}_2)|\downarrow\rangle_2 & \psi_{\mathrm{B}}(\boldsymbol{r}_2)|\downarrow\rangle_2 \end{vmatrix}. \tag{4.4.43-d}$$

States Ψ_1 and Ψ_4 are antisymmetric in the spatial variables and symmetric in the spin variable; they correspond to the components $S^z = \pm 1$ of a triplet. The $S^z = 0$ component of the triplet and the spin-singlet state – which is symmetric in the spatial variables – are respectively given by the symmetric and antisymmetric combinations of the functions Ψ_2 and Ψ_3:

$$\Psi_{\mathrm{t}} = \frac{1}{\sqrt{2}}(\Psi_2 + \Psi_3)\,, \qquad \Psi_{\mathrm{s}} = \frac{1}{\sqrt{2}}(\Psi_2 - \Psi_3)\,. \tag{4.4.44}$$

The four wavefunctions in (4.4.43) contain the combinations that are taken into account in the Heitler–London approximation. In addition to these valence configurations, one might wish to include ionic configurations as well – since they appear even in the simplest LCAO approximation. This is possible through the inclusion of the functions

$$\Psi_5 = \frac{1}{\sqrt{2}} \begin{vmatrix} \psi_{\mathrm{A}}(\boldsymbol{r}_1)|\uparrow\rangle_1 & \psi_{\mathrm{A}}(\boldsymbol{r}_1)|\downarrow\rangle_1 \\ \psi_{\mathrm{A}}(\boldsymbol{r}_2)|\uparrow\rangle_2 & \psi_{\mathrm{A}}(\boldsymbol{r}_2)|\downarrow\rangle_2 \end{vmatrix}, \tag{4.4.45-a}$$

$$\Psi_6 = \frac{1}{\sqrt{2}} \begin{vmatrix} \psi_{\mathrm{B}}(\boldsymbol{r}_1)|\uparrow\rangle_1 & \psi_{\mathrm{B}}(\boldsymbol{r}_1)|\downarrow\rangle_1 \\ \psi_{\mathrm{B}}(\boldsymbol{r}_2)|\uparrow\rangle_2 & \psi_{\mathrm{B}}(\boldsymbol{r}_2)|\downarrow\rangle_2 \end{vmatrix}. \tag{4.4.45-b}$$

The total two-particle wavefunction can be written as the linear combination of the six determinants Ψ_i,

$$\Psi = \sum_i c_i \Psi_i\,. \tag{4.4.46}$$

[12] J. C. Slater, 1929.

The Slater determinants may also be expressed in terms of the wavefunctions of the molecular orbitals rather than the atomic wavefunctions. The six functions can then be chosen as

$$\Phi_1 = \frac{1}{\sqrt{2}} \begin{vmatrix} \psi_+(\boldsymbol{r}_1)|\uparrow\rangle_1 & \psi_-(\boldsymbol{r}_1)|\uparrow\rangle_1 \\ \psi_+(\boldsymbol{r}_2)|\uparrow\rangle_2 & \psi_-(\boldsymbol{r}_2)|\uparrow\rangle_2 \end{vmatrix}, \tag{4.4.47-a}$$

$$\Phi_2 = \frac{1}{\sqrt{2}} \begin{vmatrix} \psi_+(\boldsymbol{r}_1)|\uparrow\rangle_1 & \psi_+(\boldsymbol{r}_1)|\downarrow\rangle_1 \\ \psi_+(\boldsymbol{r}_2)|\uparrow\rangle_2 & \psi_+(\boldsymbol{r}_2)|\downarrow\rangle_2 \end{vmatrix}, \tag{4.4.47-b}$$

$$\Phi_3 = \frac{1}{\sqrt{2}} \begin{vmatrix} \psi_+(\boldsymbol{r}_1)|\uparrow\rangle_1 & \psi_-(\boldsymbol{r}_1)|\downarrow\rangle_1 \\ \psi_+(\boldsymbol{r}_2)|\uparrow\rangle_2 & \psi_-(\boldsymbol{r}_2)|\downarrow\rangle_2 \end{vmatrix}, \tag{4.4.47-c}$$

$$\Phi_4 = \frac{1}{\sqrt{2}} \begin{vmatrix} \psi_+(\boldsymbol{r}_1)|\downarrow\rangle_1 & \psi_-(\boldsymbol{r}_1)|\uparrow\rangle_1 \\ \psi_+(\boldsymbol{r}_2)|\downarrow\rangle_2 & \psi_-(\boldsymbol{r}_2)|\uparrow\rangle_2 \end{vmatrix}, \tag{4.4.47-d}$$

$$\Phi_5 = \frac{1}{\sqrt{2}} \begin{vmatrix} \psi_-(\boldsymbol{r}_1)|\uparrow\rangle_1 & \psi_-(\boldsymbol{r}_1)|\downarrow\rangle_1 \\ \psi_-(\boldsymbol{r}_2)|\uparrow\rangle_2 & \psi_-(\boldsymbol{r}_2)|\downarrow\rangle_2 \end{vmatrix}, \tag{4.4.47-e}$$

$$\Phi_6 = \frac{1}{\sqrt{2}} \begin{vmatrix} \psi_+(\boldsymbol{r}_1)|\downarrow\rangle_1 & \psi_-(\boldsymbol{r}_1)|\downarrow\rangle_1 \\ \psi_+(\boldsymbol{r}_2)|\downarrow\rangle_2 & \psi_-(\boldsymbol{r}_2)|\downarrow\rangle_2 \end{vmatrix}. \tag{4.4.47-f}$$

It can be easily shown that in the LCAO approximation – where the wavefunctions of the molecular orbitals are constructed from the atomic functions in the form (4.4.34) – the functions Φ_i in (4.4.47) are linear combinations of the Ψ_i given in (4.4.43) and (4.4.45):

$$\Phi_1 = -\Psi_1\,, \qquad \Phi_6 = -\Psi_4 \tag{4.4.48}$$

and

$$\begin{aligned} \Phi_2 &= \frac{1}{\sqrt{2}}(\Psi_2 - \Psi_3 + \Psi_5 + \Psi_6)\,, \\ \Phi_3 &= \frac{1}{\sqrt{2}}(-\Psi_2 - \Psi_3 + \Psi_5 - \Psi_6)\,, \\ \Phi_4 &= \frac{1}{\sqrt{2}}(-\Psi_2 - \Psi_3 - \Psi_5 + \Psi_6)\,, \\ \Phi_5 &= \frac{1}{\sqrt{2}}(-\Psi_2 + \Psi_3 + \Psi_5 + \Psi_6)\,. \end{aligned} \tag{4.4.49}$$

In this sense the two choices are equivalent.

4.4.7 Hybridized Orbitals

Up to now we have focused on how the covalent bond is formed between *two* atoms. In solids each atom has to make bonds with several neighbors. As a generalization of the previous findings, each atom is expected to participate in as many covalent bonds as they have singly occupied electron states. The electron configuration of a ground-state carbon atom is $1s^2 2s^2 2p^2$ (see Appendix B), which means that only the two 2p electrons are on open shells. This would then imply that each carbon atom could take part in two covalent bonds. However, this is known not to be the case: carbon is tetravalent. The reason for this is that one of the 2s electrons is excited to an empty 2p state, transforming the configuration into $1s^2 2s^1 2p^3$, and thus making further bonds possible. Although in a free atom this configuration has higher energy than the ground state, the difference is compensated for by the decrease in energy caused by the appearance of new bonds.

Wavefunctions are customarily expressed in terms of spherical harmonics:

$$\psi_{nlm}(\boldsymbol{r}) = R_{nl}(r) Y_l^m(\theta, \varphi)\,. \tag{4.4.50}$$

However, for p ($l = 1$) states the real linear combinations ψ_{p_x}, ψ_{p_y}, and ψ_{p_z} offer a more practical choice:

$$\begin{aligned}\psi_{np_\alpha}(\boldsymbol{r}) &= R_{n1}(r) \begin{cases} \dfrac{-1}{\sqrt{2}} \left(Y_1^1(\theta,\varphi) - Y_1^{-1}(\theta,\varphi)\right) \\ \dfrac{\mathrm{i}}{\sqrt{2}} \left(Y_1^1(\theta,\varphi) + Y_1^{-1}(\theta,\varphi)\right) \\ Y_1^0(\theta,\varphi) \end{cases} \\ &= \left(\frac{3}{4\pi}\right)^{1/2} R_{n1}(r) \begin{cases} x/r\,, \\ y/r\,, \\ z/r\,. \end{cases} \end{aligned} \tag{4.4.51}$$

These formulas give high electron densities along the coordinate axes, as shown in Fig. 4.10.

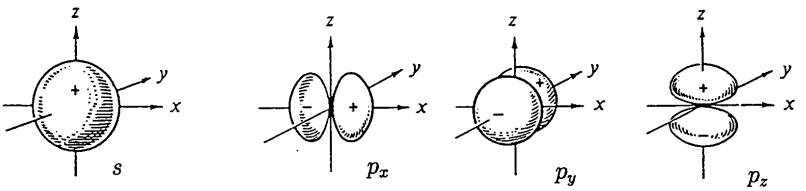


Fig. 4.10. Sketch of the wavefunctions of states s, p_x, p_y, and p_z

To get a better insight into the states 2s, $2p_x$, $2p_y$, and $2p_z$, consider the four orthogonal combinations

$$\begin{aligned}
\psi_1 &= \tfrac{1}{2}[\psi_{2s} + \psi_{2p_x} + \psi_{2p_y} + \psi_{2p_z}]\,, \\
\psi_2 &= \tfrac{1}{2}[\psi_{2s} + \psi_{2p_x} - \psi_{2p_y} - \psi_{2p_z}]\,, \\
\psi_3 &= \tfrac{1}{2}[\psi_{2s} - \psi_{2p_x} + \psi_{2p_y} - \psi_{2p_z}]\,, \\
\psi_4 &= \tfrac{1}{2}[\psi_{2s} - \psi_{2p_x} - \psi_{2p_y} + \psi_{2p_z}]\,.
\end{aligned} \tag{4.4.52}$$

From the formulas for the wavefunctions of the s- and p-states it is straightforward to show that these sp^3 hybrid functions give high electron densities in the directions of the four vertices of a regular tetrahedron. This is illustrated in Fig. 4.11.

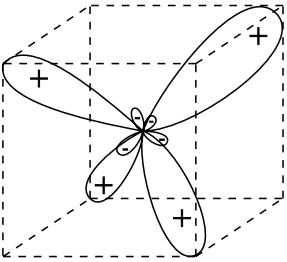


Fig. 4.11. Spatial distribution of the electron density in the states that correspond to the four sp^3 hybrid wavefunctions

Now consider two carbon atoms, in $\boldsymbol{R}_{\mathrm{A}}$ and $\boldsymbol{R}_{\mathrm{B}}$. Denoting the corresponding hybrid one-particle wavefunctions by $\psi_i^{(\mathrm{A})}$ and $\psi_j^{(\mathrm{B})}$, one can construct two-particle wavefunctions along the same lines as in (4.4.23):

$$\begin{aligned}
\psi_{ij}(\boldsymbol{r}_1, \boldsymbol{r}_2) = N \Big[&\psi_i^{(\mathrm{A})}(\boldsymbol{r}_1)\psi_j^{(\mathrm{B})}(\boldsymbol{r}_2) + \psi_j^{(\mathrm{B})}(\boldsymbol{r}_1)\psi_i^{(\mathrm{A})}(\boldsymbol{r}_2) \\
&+ \lambda\psi_i^{(\mathrm{A})}(\boldsymbol{r}_1)\psi_i^{(\mathrm{A})}(\boldsymbol{r}_2) + \lambda\psi_j^{(\mathrm{B})}(\boldsymbol{r}_1)\psi_j^{(\mathrm{B})}(\boldsymbol{r}_2) \Big]\,.
\end{aligned} \tag{4.4.53}$$

Among the numerous such possibilities there are only a few with low-energy covalent bonds: those for which the hybrid wavefunctions of two neighboring atoms display high electron density along the line joining the two atoms. The covalent bond is therefore highly directional. This approach is well suited for the description of bonds in diamond (built up of carbon atoms) and in semiconducting materials with a similar structure.

Besides the hybrid sp^3 wavefunctions presented above, other hybrid states are possible as well. For example, one s- and one p-state can give rise to an sp state; one s- and two p-states to an sp^2 state. p- and d-states can also lead to hybrid orbitals. Needless to say, the bonds are now oriented in other directions

than the vertices of a tetrahedron. We shall revisit this question at the end of Section 7.6 on the relationship between crystal structure and bonding.

4.4.8 Covalent Bonds in Solids

Covalent solids are held together by networks of directional covalent bonds between neighboring atoms, where the wavefunctions of covalent bonds are constructed from pairs of atomic states. If each bonding state is occupied by two electrons of opposite spins – i.e., each bond is saturated – then the electrons that participate in bonding are localized in space and cannot contribute to electrical conductivity. That is why covalent materials are usually insulators or semiconductors. This is the case for diamond and for two further elements of the carbon group: germanium and silicon, which have a diamond structure. Here each atom has four sp^3 hybrid states, and the bonds formed by them make up a tetrahedral network, as shown in Fig. 7.16(a). The same structure is seen in semiconducting compounds formed by elements in groups 13 (IIIA) and 15 (VA) of the periodic table.

As it was mentioned in the previous subsection, besides sp^3 wavefunctions, other hybrid states may also give rise to covalent bonding – however, their spatial directionality depends on which states are hybridized. The orientation of the bonds plays a crucial role in determining the crystal structure. The development of short-range order in the amorphous state of covalently bonded solids is also related to the directionality of the bonds.

The cohesive energy of covalent crystals is given by the sum of the binding energies of individual bonds. The binding energies of some typical covalent bonds are listed in Table 4.5. Much larger than their counterparts for molecular crystals, these values are comparable to the energies of ionic crystals.

Table 4.5. Binding energy of some typical covalent bonds (in units of eV and kJ/mol)

Bond	eV	kJ/mol	Bond	eV	kJ/mol
H–H	4.48	432	C–H	4.28	413
N–N	1.65	159	C–N	3.16	305
P–P	2.08	201	N–H	4.03	389
C–C	3.58	346	Al–P	2.13	205
Si–Si	2.30	222	Si–C	3.17	306
Ge–Ge	1.95	188	Ga–As	1.63	157
O–O	1.47	142	Ga–P	1.78	172

4.5 Metallic Bond

A great part of the chemical elements have fewer electrons on the incomplete shells than what is necessary for having saturated covalent bonds between each pair of neighboring atoms in the solid state. Electrons participating in unsaturated bonds are not localized. One may also say that the atoms lose these outermost (in general, s or p) electrons. While the positively charged ions left behind are arranged in a more or less regular pattern, the freed electrons fill the region among the ions almost evenly. This moving cloud of electrons gives rise to *metallic bonding*. In transition metals, where incomplete d-shells are also found under the outermost shell, further electrons may participate in metallic bonding. The same kind of bonding may appear in materials built up of molecules with incomplete shells.

The wavefunction of such an electron system cannot be written as the product of the wavefunctions of pairs of electrons forming bonds. It has to be chosen in such a way that it should show explicitly the antisymmetry with respect to the interchange of the coordinates of any two electrons. This can be done through the generalization of the formulas in (4.4.43) or (4.4.47), using functions of the Slater determinant form. The analysis of such systems, the determination of electronic energies and states, and the study of those properties of solids that are due to electrons will be among our most important tasks. Volumes 2 and 3 are devoted almost exclusively to these issues. To a large extent, the present volume serves as preparation for this.

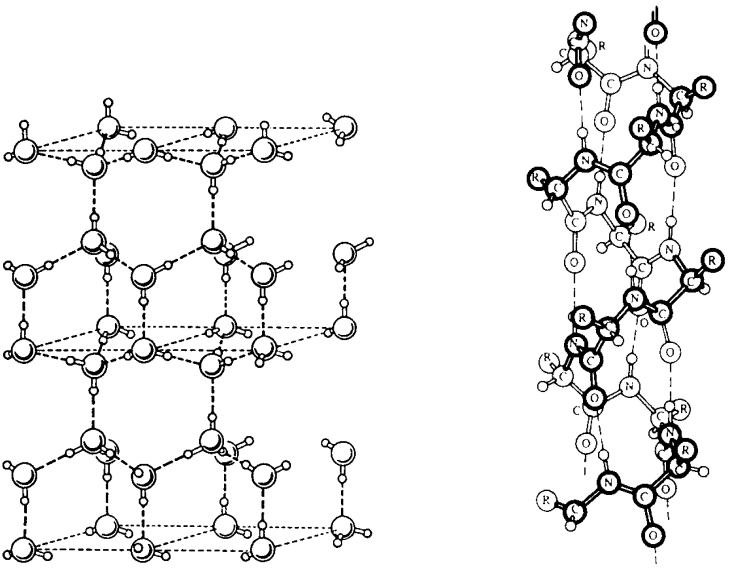
The determination of the total energy of metals is a difficult problem of solid-state physics. We shall take a closer look at it in Chapter 30, after the study of electron states. Here we only mention that the cohesive energy per atom is usually 1–5 eV.

4.6 The Hydrogen Bond

The hydrogen atom has a single electron on the 1s shell – lacking another one to have the shell closed. That is why covalent bonding would permit hydrogen to be linked to a single other ion. Due to its small size, even with ionic bonding only two ions can be tightly packed around the proton. However, because of its high ionization energy (13.6 eV) hydrogen is not easily ionized. Instead of participating in such bonds, hydrogen can create a special bond between highly electronegative atoms like fluorine, oxygen, or nitrogen. This is the *hydrogen bond* (or, as it is called in several languages, the "hydrogen bridge").

In this type of bonding the hydrogen atom is not located at the midpoint between the F, O, N atoms, but has two symmetrical equilibrium positions between which it oscillates.

Crystalline ice is held together by such bonds, as shown in Fig. 4.12(a) – but hydrogen bonds play an important role in water, too. The distance between oxygen atoms is about 2.75 Å in the ground state of ice and about 2.9 Å



(*a*) (*b*)

Fig. 4.12. Hydrogen bonds in (*a*) ice; and (*b*) the alpha helix structure of a polypeptide

in water. The two equilibrium positions of the hydrogen atom are located at a distance of about 1 Å from either of them. The bond between the hydrogen and the nearby oxygen can be considered as covalent, although the huge difference in electronegativity makes it highly polar. That is why the hydrogen can be linked to another oxygen atom on the other side, at a distance of 1.8–2 Å. The hydrogen bond is stronger on one side and weaker on the other; this is reflected in the notation O–H$\cdots$O. The bond is relatively weak; breaking it requires some 0.2 eV of energy. Between two fluorine atoms the binding energy is 0.29 eV.

In crystalline ice each oxygen atom is connected to four others through hydrogen bonds. If a snapshot were made of the structure, in two of these bonds the hydrogen would be in its equilibrium position closer to the oxygen, and in the farther one in the other two. It can be shown that there are exponentially many states that satisfy this condition. This high degeneracy exists in the ground state, too – giving rise to a finite entropy at zero temperature. That is why ice is a model of choice in statistical physics.

Even more important is the fact that the structure of proteins is determined by hydrogen bonds (with binding energies of 8 to 40 kJ/mol) between the CO and NH groups of the polypeptide chains. The blueprint of life, the double helix of DNA is also held together by hydrogen bonds between base pairs. The schematic structure of an alpha helix is shown in Fig. 4.12(*b*).

Hydrogen bonding is therefore essential for biological materials – however, substances held together by hydrogen bonds do not play an important role in solid-state physics. That is why we shall not pursue the study of this type of bond any further.

Further Reading

1. J. K. Burdett, *Chemical Bonding in Solids*, Oxford University Press, Oxford (1995).
2. R. McWeeny, *Coulson's Valence*, Third Edition, Oxford University Press, Oxford (1979).
3. J. N. Murrell, S. F. A. Kettle, and J. M. Tedder, *The Chemical Bond*, Second Edition, John Wiley and Sons Ltd., Chichester (1985).
4. L. C. Pauling, *The Nature of the Chemical Bond and the Structure of Molecules and Crystals: An Introduction to Modern Structural Chemistry*, 3rd edition, Cornell University Press, Ithaka, N.Y. (1960).
5. D. Pettifor, *Bonding and Structure of Molecules and Solids*, Clarendon Press, Oxford (1995).
6. G. S. Rohrer, *Structure and Bonding in Crystalline Materials*, Cambridge University Press, Cambridge (2001).

5

Symmetries of Crystals

It has already been mentioned that crystalline materials constitute a highly important group of solids. In what follows, we shall be mostly concerned with the study of the properties of such materials. The reasons for this are not purely historical. On the one hand, crystalline materials are used in many applications. On the other hand, quite a few electronic properties do not depend on structure, and the theoretical study of the behavior of electrons within solids is much simpler in crystalline materials than in noncrystalline ones.

One of the most striking features of natural crystals is that they are usually bounded by regular plane faces. The regularity of the form – and thus the rotations and reflections that take the crystal into itself – are even better seen on artificially grown crystals.[1] As R.-J. Haüy pointed out already in 1784, the regular form of carefully grown crystals implies that crystals are built up of regularly arranged small blocks. The regularity of the overall form is therefore the consequence of an internal regularity. It was well after Haüy's time that atoms and ions were identified as the elementary building blocks, and it was shown that in the crystalline state atoms (or ions) with identical surroundings are arranged in a regular periodic array – implying that the crystal exhibits long-range order.

In the present chapter we shall discuss the geometrical characteristics of crystalline solids. Although atoms are not arranged in perfect order in real crystals – which may have important consequences on the properties of the solid –, here we shall assume that the crystal is ideal, i.e., perfectly regular and without defects. The various types of defects will be presented in Chapter 9. Noncrystalline solid structures will be discussed in Chapter 10.

[1] Being highly malleable, the crystalline symmetry of metals is almost always disguised in finished products.

5.1 Translational Symmetry in Crystals

The characteristic regularity of crystalline structures is usually formulated in terms of the invariance of the crystal under certain discrete translations. As we saw in Chapter 2, there exist mesomorphic phases in which the material possesses discrete translational symmetries along one or two directions only and continuous ones along the others. Below we shall consider solids which are taken into themselves only by discrete translations in all three spatial directions. To put it more rigorously: A solid is said to be crystalline if there exist three non-coplanar vectors $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, and $\boldsymbol{a}_3$ such that for any three integers n_1, n_2, and n_3 translation through the vector

$$\boldsymbol{t}_n = n_1\boldsymbol{a}_1 + n_2\boldsymbol{a}_2 + n_3\boldsymbol{a}_3 \tag{5.1.1}$$

takes the sample into itself – while this is not the case for linear combinations in which the coefficients are not all integers. The vectors $\boldsymbol{a}_i$ appearing in the definition are called *primitive translation vectors* or *primitive vectors*, and are said to *span* or *generate* the lattice. This formulation of discrete translational symmetry is the mathematical definition of ideal crystals.

Mathematically speaking, these translations form a group, T_3. For convenience, we shall frequently refer to two-dimensional crystals. These are planar arrangements that are translationally invariant in the two planar dimensions (under the elements of the group T_2).

Invariance under discrete translations does not apply solely to the regular array of spatially localized atoms. It also implies that the charge density $\rho_{\mathrm{e}}(\boldsymbol{r})$ of delocalized electrons, as well as the internal degrees of freedom (e.g., spin density) must be left unaltered by a translation through $\boldsymbol{t}_n$. Consequently, for any local observable quantity $F(\boldsymbol{r})$

$$F(\boldsymbol{r}) = F(\boldsymbol{r} + \boldsymbol{t}_n) \tag{5.1.2}$$

must hold. For nonobservable quantities – such as electron wavefunctions – this invariance does not apply. As we shall later, when considered separately, localized atoms and the charge density of delocalized electrons may exhibit different translational symmetries. Then the true translational symmetries of the whole system are translations through those vectors $\boldsymbol{t}_n$ that transform both the localized atoms and the charge density into themselves.

5.1.1 Translational Symmetry in Finite Crystals

Strictly speaking, the previous formulation is valid only for infinite crystals. Finite crystals obviously cannot be taken perfectly into themselves by translations. However, where the original and translated samples overlap, the atomic arrangements must be identical – with the possible exception of a few atomic layers near the surface. The previous statement can be reformulated to be valid for finite samples: *A sample of finite extent is said to be crystalline if*

it is built up of identical elementary blocks (atoms, molecules, or groups of atoms/molecules) whose positions are given by the vectors in (5.1.1), and the triplets of integers n_1, n_2, n_3 form a compact region. Figure 5.1 shows such a finite crystalline structure built up of groups of five atoms.

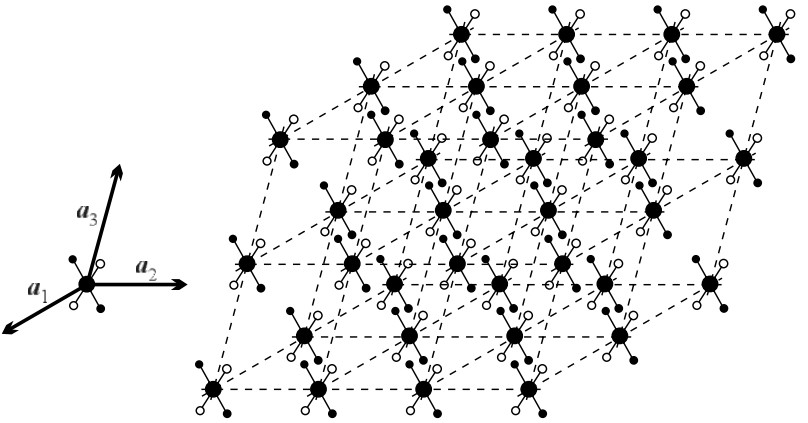


Fig. 5.1. A finite crystalline structure: a finite regular three-dimensional array of identical groups of atoms

A similar situation is encountered in crystal-growth processes. First, a nucleus is formed at random, and then other atoms/molecules condense onto it in a regular pattern. This construction of the crystal implies that whichever atom (of position vector $\boldsymbol{r}$) of the starting group of atoms is chosen, identical atoms (in equivalent surroundings) are found at all points

$$\boldsymbol{r}' = \boldsymbol{r} + \boldsymbol{t}_n = \boldsymbol{r} + n_1\boldsymbol{a}_1 + n_2\boldsymbol{a}_2 + n_3\boldsymbol{a}_3\,, \tag{5.1.3}$$

where the coefficients n_i are integers, just as in (5.1.1). In general, when $\boldsymbol{r}$ and $\boldsymbol{r}'$ are related through (5.1.3), $F(\boldsymbol{r}') = F(\boldsymbol{r})$ holds for the spatial distribution of electron density and any local observable physical quantity $F(\boldsymbol{r})$ – with the possible exception of the vicinity of the surface. Thus for finite crystals the relation (5.1.2) is valid only for positions $\boldsymbol{r}$ and $\boldsymbol{r} + \boldsymbol{t}_n$ that are both inside the sample.

5.1.2 The Choice of Primitive Vectors

One usually faces the opposite problem as in the previous theoretical construction of crystals. Given some regular arrangement of the atoms, one needs to determine the primitive vectors $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, and $\boldsymbol{a}_3$. The choice of these vectors is ambiguous: if a vector triplet $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$ is found such that the crystal is invariant under the translations $\boldsymbol{t}_n$ given in (5.1.1), then the same invariance holds for any other vector triplet $\boldsymbol{a}'_1$, $\boldsymbol{a}'_2$, $\boldsymbol{a}'_3$ that is a linearly independent set

of linear combinations of the previous one with integer coefficients. However, the new vector triplet is not necessarily primitive (even when the old one is): it does not necessarily satisfy the condition that all translation vectors should be given by integer coefficients.

A possible choice for the primitive vectors $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ and a new set $\boldsymbol{a}'_1$, $\boldsymbol{a}'_2$ are shown in Fig. 5.2 for a two-dimensional crystal. With either choice, translations through $\boldsymbol{t}_n = n_1\boldsymbol{a}_1 + n_2\boldsymbol{a}_2$ or $\boldsymbol{t}'_n = n'_1\boldsymbol{a}'_1 + n'_2\boldsymbol{a}'_2$ take the lattice into itself: each point is moved to an equivalent one. It is immediately seen that the two choices are nevertheless inequivalent: when translations through integer n'_is only are allowed, the choice on the right does not cover all the points that are equivalent to $\boldsymbol{r}$: $\boldsymbol{a}'_1$ and $\boldsymbol{a}'_2$ are thus not primitive vectors.

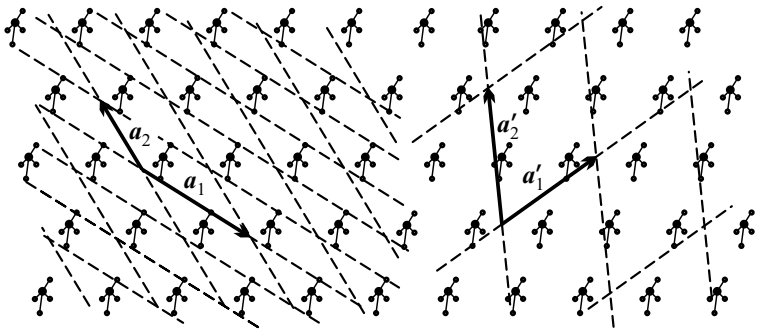


Fig. 5.2. A possible choice for the primitive vectors and a pair of new vectors obtained through their linear combination in a two-dimensional crystal

Therefore the requirement that the vectors $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$ be primitive – in other words: that the translations $\boldsymbol{t}_n$ obtained with triplets of integers n_1, n_2, n_3 yield all points $\boldsymbol{r}' = \boldsymbol{r} + \boldsymbol{t}_n$ in the crystal that are equivalent to $\boldsymbol{r}$ – is tantamount to stipulating that the volume of the parallelepiped spanned by the non-coplanar vectors $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$ be minimum.

The minimum volume requirement is not perfectly restrictive either: it allows for many possible choices of the primitive vector, as shown in Fig. 5.3. For clarity, instead of the entire crystal only points that are equivalent to a selected one are marked.

Therefore it is useful to impose the restriction that the length of the vectors should be the smallest possible. Among the possible choices shown in Fig. 5.3 the leftmost meets this requirement. In many cases it is possible to go even further. Later, when the rotational symmetry of crystals will have been presented, it will be appreciated that primitive vectors should possibly be chosen in such a way that they are related through some symmetry operation. Moreover, in certain cases even the requirements of minimum volume and minimum length are abandoned so that rotational symmetry should be manifest. The lattice is then characterized by vectors $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$ that are

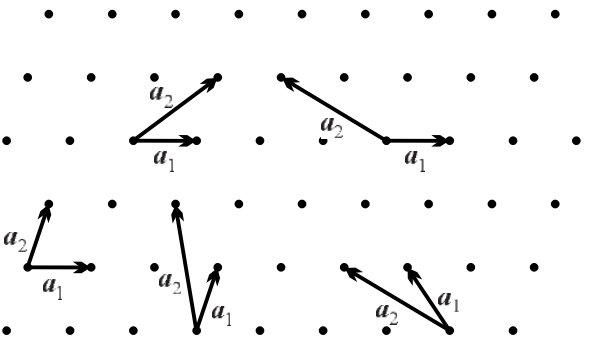


Fig. 5.3. Several possible choices for the pair of primitive vectors in a two-dimensional crystal. In each case every equivalent point can be reached by a integer linear combination of the primitive vectors

more symmetric than the previously used primitive vectors $\boldsymbol{a}_1$, $\boldsymbol{a}_2$ and $\boldsymbol{a}_3$. Since $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$ are not primitive vectors, a part of the equivalent sites in the original lattice can be reached only via triplets of coefficients that contain nonintegers.

5.1.3 Bravais Lattice and Basis

Once the primitive vectors have been chosen and an origin selected, the endpoints of the vectors (5.1.1) marked for all possible integer values of n_1, n_2, and n_3 make up a regular albeit empty lattice. This is called the *point lattice* or *Bravais lattice* of the crystal – named after the French crystallographer A. Bravais who was the first to determine correctly the possible lattice types in 1850. The points of a Bravais lattice are thus given by

$$\boldsymbol{R}_n = n_1\boldsymbol{a}_1 + n_2\boldsymbol{a}_2 + n_3\boldsymbol{a}_3\,. \tag{5.1.4}$$

This will be important when we shall try to exploit the consequences of rotational symmetry.

Note that the lattice concept introduced above is different from that used in statistical physics. Statistical physical models are often solved on honeycomb and kagome (basketweave)[2] lattices (shown in Fig. 5.4). However, their vertices (sites) do not make up a lattice in the above sense. There are no primitive vectors whose linear combinations with integer coefficients give all the vertices and only the vertices. This is why when we present all possible types of planar lattices below these structures will not appear.

The knowledge of the primitive vectors, i.e., the point lattice is not sufficient for characterizing the crystal structure – even when only the spatial

[2] The name comes from the Japanese word kagome, meaning the pattern of holes ("me" = holes, literally "eyes") in a woven (bamboo) basket ("kago").

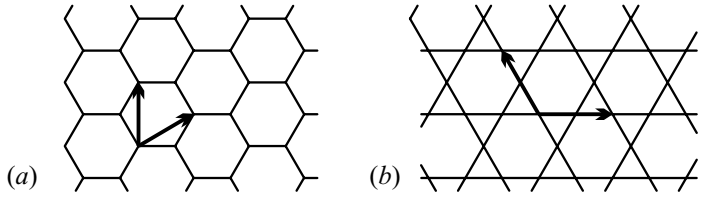


Fig. 5.4. "Lattices" that are not Bravais lattices, and their primitive vectors: (*a*) honeycomb lattice; (*b*) kagome lattice

arrangement of atoms, i.e., the geometry of the crystal is considered. One must also identify the group of atoms whose regularly repeated pattern makes up the lattice. This group of atoms is called the *basis* or *motif* of the crystal. The honeycomb net and the kagome lattice are thus Bravais lattices with a two- and a three-point basis, respectively.

When the same Bravais lattice is decorated with different bases, different crystal structures are obtained. Conversely, crystals with strikingly different structure may share the same underlying Bravais lattice. It is easily seen that the two-dimensional crystals shown in parts (*a*) and (*b*) of Fig. 5.5 have different bases but a common Bravais lattice, which is shown in part (*c*). The crystal structure is therefore determined by its Bravais lattice *and* its basis.

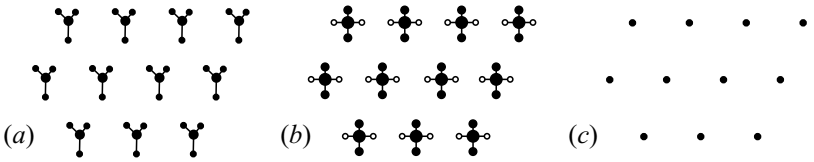


Fig. 5.5. (*a*) and (*b*): Two-dimensional crystals that are built up of different groups of atoms but exhibit the same translational symmetries. (*c*): Their common Bravais lattice

When the basis consists of a single atom, then the origin of the coordinate system can always be chosen so that each atom should occupy a lattice site $\boldsymbol{R}_n$. On the other hand, in crystals with a p-point basis the arrangement of the atoms in the basis should appear exactly the same around each lattice site; therefore atomic positions are given by the sum of a lattice vector $\boldsymbol{R}_n$ and a vector $\boldsymbol{r}_\mu$ ($\mu = 1, 2, \ldots, p$), which is referred to the corresponding lattice site, thus lies in the unit cell.

5.1.4 Primitive Cells, Wigner–Seitz Cells, and Bravais Cells

The basis is not always pictured as a group of several point-like atoms. It can also be thought of as an extended charge distribution. It is then not

just a collection of atoms but an elementary region with atoms (ions) and a smeared-out electron cloud. Such an elementary region called *primitive unit cell* or *primitive cell* contains exactly one lattice point. When it is repeated according to a regular spatial pattern, i.e., translated through all vectors $\boldsymbol{t}_n$, the volume of the crystal is covered in its entirety. This condition does not impose a strong restriction on the shape of the primitive cell. It does not need to possess high symmetry, nor are its faces required to be plane (or in the two-dimensional case: its sides to be straight). Complete filling of the space or complete tiling of the plane requires only that opposite faces (sides) could be moved into one another by simple translations – leaving considerable freedom in the choice of the figures. Beautiful examples for periodic tilings of the plane are found on embroideries, wall decorations, and mosaics. In his artistic program the Dutch artist M. C. Escher[3] illustrated possible symmetries in the complete tiling of the plane using human and animal figures. Some of his works are shown in Fig. 5.6, as well as in Figs. 5.27 and 5.31 in the later parts of this chapter.

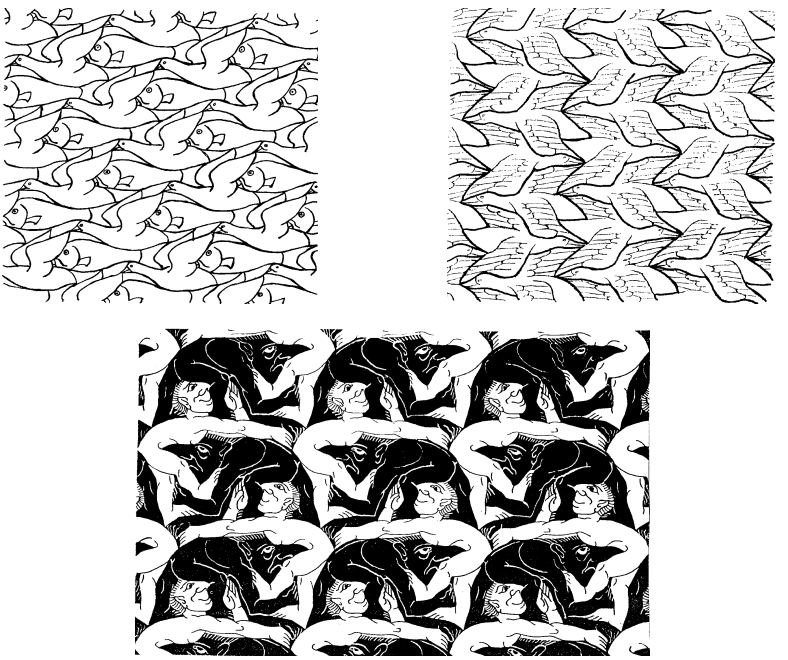


Fig. 5.6. Escher's drawings show the complete tiling of the plane with human and animal figures

[3] M. C. Escher, 1898–1972.

It is a straightforward matter to prove that it is often impossible to fill the space or tile the plane with patterns of high symmetry. The plane cannot be perfectly tiled with identical circles, and space cannot be entirely filled with identical spheres. For mathematical simplicity, the electron states are nevertheless often determined within a sphere centered on a lattice site (the *Wigner–Seitz sphere*). This approximation and its problems will be discussed in Chapter 19 of Volume 2 where the study of electron states is presented.

It should also be mentioned here – without going into the details of the geometrical problem of finding those solids bounded by plane faces that can fill up space perfectly – that perfect tilings of the plane are possible with regular or distorted triangles, quadrilaterals, or hexagons while the tiling will be imperfect when regular pentagons or octagons are used, as shown in Fig. 5.7. In the case of pentagons allowing for nonidentical orientation does not help either. It will be shown in Section 10.2 on quasicrystals that the plane can be tiled in such a manner that local fivefold symmetry is manifest – provided that the requirement of complete translational symmetry is relaxed and only that of quasiperiodicity is imposed.

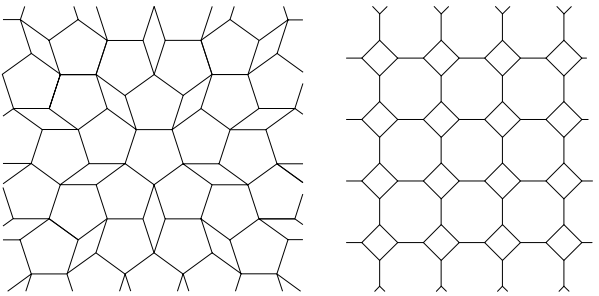


Fig. 5.7. Incomplete tiling of the plane with regular pentagons and octagons

A common choice for the primitive cell is the parallelepiped spanned by the primitive vectors $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, and $\boldsymbol{a}_3$. Clearly, a regular array of such parallelepipeds fills the space completely, without overlaps. Figure 5.8(a) shows the division of the two-dimensional periodic crystal structure of Fig. 5.2 into cells spanned by the minimum-length primitive vectors $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$.

The freedom in the choice of the origin is customarily abandoned in favor of choosing a characteristic atom of the crystal as the starting point of the primitive vectors. In this construction the group of atoms around a particular lattice site is usually divided among several primitive cells. Another disadvantage of this choice for the primitive cell is that translational symmetry alone is taken into account: the internal structure and symmetries of the figure the crystal is built up of are completely ignored. That is why in many cases – in particular, when the electron states are to be determined – the crystal is divided into primitive cells in such a manner that in addition to the atoms,

the core electrons should also belong to the same region. Such a division is obtained through *Dirichlet's construction*.

Consider the points of a crystal lattice. In Dirichlet's construction one associates with each lattice point the region of space that is closer to that point than to any other lattice point. To construct these regions, vectors are drawn from the point in question to its first, second and perhaps third neighbors, and then perpendicular planes are drawn through the midpoint of each vector. The primitive cell obtained this way is called the *Wigner–Seitz cell*.[4] Part (*c*) of Fig. 5.8 shows the tessellation of the two-dimensional crystal of parts (*a*) and (*b*) into Wigner–Seitz cells.

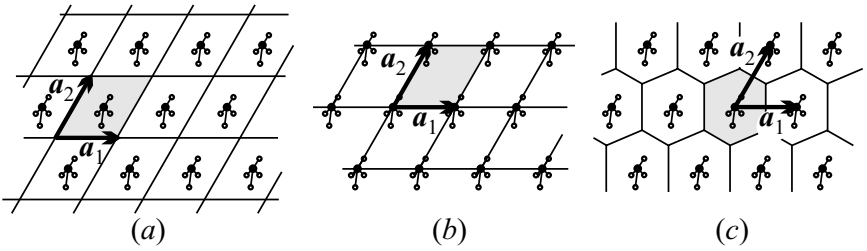


Fig. 5.8. Three different choices for the primitive cells (shaded areas) that allow for perfect tiling of the two-dimensional crystal. (*a*) and (*b*) differ only in the choice of the origin; (*c*) tessellation by Wigner–Seitz cells

It has been mentioned that the relative orientation of the primitive vectors does not necessarily reflect the crystal symmetries, which is why sometimes longer but more symmetric vectors, $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$ are used instead. As an example, consider the plane lattice shown in Fig. 5.9. The primitive vectors are of different lengths, and their angle is such that $2\boldsymbol{a}_2 - \boldsymbol{a}_1$ is perpendicular to $\boldsymbol{a}_1$.

Instead of $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ one might choose the vectors $\boldsymbol{a} = \boldsymbol{a}_1$ and $\boldsymbol{b} = 2\boldsymbol{a}_2 - \boldsymbol{a}_1$. This choice gives a rectangular (nonprimitive) unit cell instead of the rhomboid-shaped primitive cell. Horizontal and vertical reflection symmetry – which are the consequence of the particular relation between $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ ($2\boldsymbol{a}_2 - \boldsymbol{a}_1 \perp \boldsymbol{a}_1$) as well as the appropriate choice of $\boldsymbol{b}$ – are manifest in this case. The parallelepiped spanned by these vectors is called the *conventional unit cell* or the *Bravais cell* of the crystal. The general definition of the Bravais cell will be given on page 144.

Since it may contain several crystallographically equivalent sites, the Bravais cell may have a larger volume than the primitive cell or the Wigner–Seitz cell. In the above example lattice points are found inside the rectangles as well

[4] E. Wigner and F. Seitz, 1933. When Dirichlet's construction is applied to sets of discrete points that do not necessarily form a Bravais lattice, the obtained irregular units are known as *Voronoi polyhedra* (G. Voronoi, 1908).

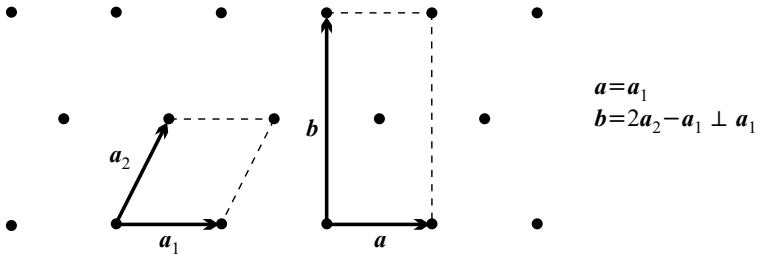


Fig. 5.9. The rhomboid-shaped primitive cell and the rectangular Bravais cell of a special oblique lattice

as at their vertices. The Bravais cell is thus not necessarily a primitive cell – it is nevertheless often a convenient choice because of its higher symmetry.

5.1.5 Crystallographic Positions, Directions, and Planes

As it was mentioned on page 114, the position vectors $\boldsymbol{r}_i$ of the atoms of the basis are referred to the points of the Bravais lattice – and these vectors do not depend on the choice of the particular cell. The position vectors are customarily expressed in terms of the primitive vectors (which are generally not mutually orthogonal):

$$\boldsymbol{r}_i = x_{i1}\boldsymbol{a}_1 + x_{i2}\boldsymbol{a}_2 + x_{i3}\boldsymbol{a}_3\,, \tag{5.1.5}$$

where the coordinates x_{i1}, x_{i2}, and x_{i3} range between 0 and 1 or $-\frac{1}{2}$ and $\frac{1}{2}$. The position vector is concisely denoted by the three coordinates $x_{i1}\,x_{i2}\,x_{i3}$. Thus $\frac{1}{2}\frac{1}{2}0$ refers to the position given by the vector $\frac{1}{2}\boldsymbol{a}_1 + \frac{1}{2}\boldsymbol{a}_2$. Negative numbers are denoted by overlines, e.g., the concise notation for the point at $\frac{1}{2}\boldsymbol{a}_1 - \frac{1}{4}\boldsymbol{a}_2$ is $\frac{1}{2}\overline{\frac{1}{4}}0$.

A specified direction in the crystal is denoted by a triplet of integers in square brackets: $[u\,v\,w]$ stands for the direction of the vector $u\,\boldsymbol{a}_1 + v\,\boldsymbol{a}_2 + w\,\boldsymbol{a}_3$. In a cubic lattice – made up of cubic primitive cells – [100], [110], and [111] denote the direction of an edge, a face diagonal, and a space diagonal. The directions of the three cube edges ([100], [010], [001]) and their opposites ($[\bar{1}00]$, $[0\bar{1}0]$, $[00\bar{1}]$) are obviously equivalent. They are collectively referred to as the $\langle 100\rangle$ directions.

In what follows, coplanar atoms will be of particular importance. Consider any three noncollinear points of a Bravais lattice and their common plane. Translational symmetry of the lattice implies that there are an infinite number of lattice points on this plane – moreover the family of equidistant parallel planes contains all the lattice points. Figure 5.10 shows this for a two-dimensional section of the lattice with three possible choices for the family of planes.

Crystal planes are also denoted by triplets of integers. To specify the three indices, consider the member of the family of planes that is closest to the plane

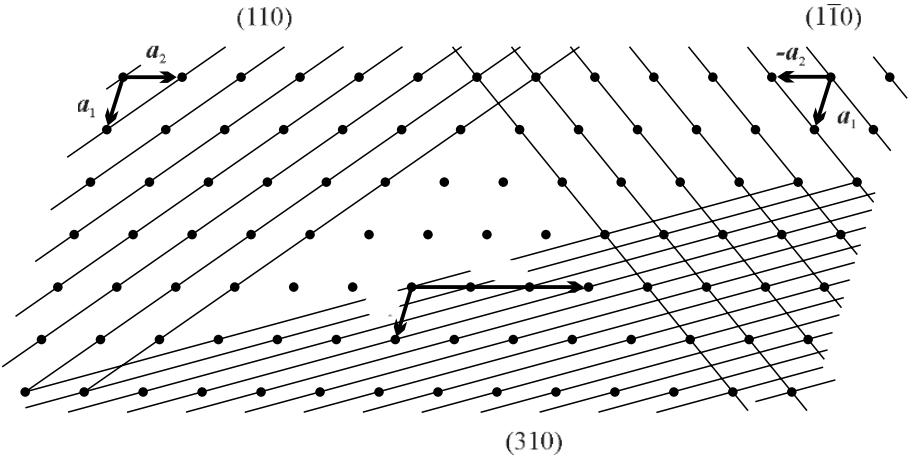


Fig. 5.10. Sections of planes through the sites of a Bravais lattice for different choices of the Miller indices (hkl)

through the origin and intercepts each of the three crystallographic axes (the lines along the primitive translation vectors) at lattice points. Using the lattice constants a_1, a_2, a_3 (i.e., the lengths of the primitive vectors) as units, these intercepts are written as $p\,0\,0$, $0\,q\,0$, and $0\,0\,r$, where p, q, and r are integers. The plane – and the family of planes as well – is then unambiguously specified by the three numbers (pqr), the *Weiss indices*.[5] It is easily shown that in the oblique coordinate system spanned by the primitive vectors the equation of the family of planes is

$$\frac{x_1}{p\,a_1} + \frac{x_2}{q\,a_2} + \frac{x_3}{r\,a_3} = n\,, \tag{5.1.6}$$

where n is an integer.

For reasons that will become apparent later, the reciprocal of these numbers – or more precisely, the smallest triplet of integers h, k, l for which

$$h : k : l = \frac{1}{p} : \frac{1}{q} : \frac{1}{r} \tag{5.1.7}$$

is used to specify the orientation of the plane. These numbers are called the *Miller indices*[6] of the plane, and are customarily given in parentheses: (hkl). Figure 5.11 shows some planes of a cubic lattice with their Miller indices.

Similarly to directions, some of the planes are also equivalent for symmetry reasons – e.g., planes (110), $(\bar{1}10)$, $(01\bar{1})$ etc. in a cubic crystal. Such equivalent planes are collectively denoted as $\{110\}$.

[5] Ch. S. Weiss, 1818.

[6] W. H. Miller, 1839. In trigonal and hexagonal structures the four Miller–Bravais indices are also commonly used.

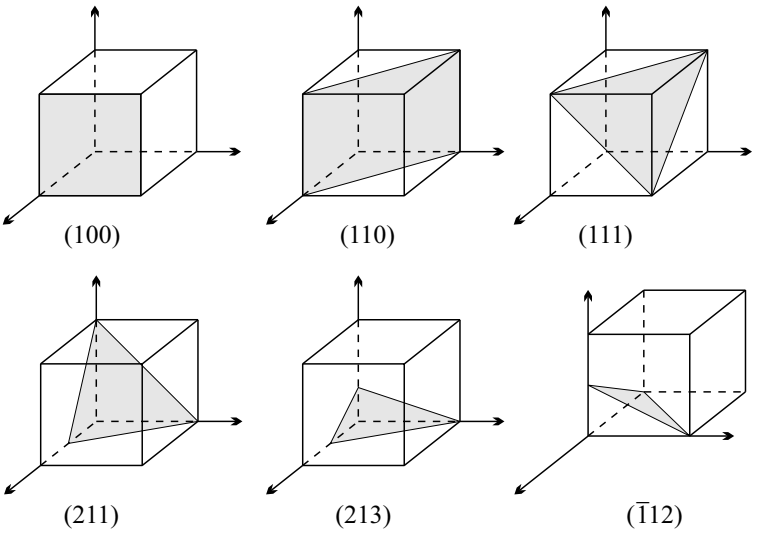


Fig. 5.11. Planes in a cubic lattice and their Miller indices

Figure 5.10 also shows that the larger the Miller indices h, k, l, the smaller the separation d_{hkl} of the planes – and the larger the distance between the nearest coplanar lattice points. Since the separation of the planes plays an important role in the experimental studies of the crystal structure, methods for its determination will be discussed later.

5.2 The Reciprocal Lattice

The structure of the crystal and the coordinates of the atoms within a primitive cell are specified in terms of the primitive vectors $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$ or the edge vectors $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ of the Bravais cell (conventional unit cell). Experimental determination of the structure and especially specification of the states within the crystal are highly simplified by using another type of lattice, the *reciprocal lattice.*[7] For future reference, below we define and present some properties of the reciprocal lattice.

5.2.1 Definition of the Reciprocal Lattice

As we have seen, the primitive vectors $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$ usually constitute an oblique system. In terms of the mutually perpendicular unit vectors $\hat{\boldsymbol{x}}$, $\hat{\boldsymbol{y}}$, $\hat{\boldsymbol{z}}$

$$\boldsymbol{a}_i = a_{ix}\hat{\boldsymbol{x}} + a_{iy}\hat{\boldsymbol{y}} + a_{iz}\hat{\boldsymbol{z}}\,, \qquad i = 1, 2, 3\,. \tag{5.2.1}$$

[7] P. P. Ewald, 1913.

When the coordinates of the primitive vectors are written as a column vector

$$\boldsymbol{a}_i = \begin{pmatrix} a_{ix} \\ a_{iy} \\ a_{iz} \end{pmatrix} \tag{5.2.2}$$

and then as a matrix

$$A = \begin{pmatrix} a_{1x}\ a_{2x}\ a_{3x} \\ a_{1y}\ a_{2y}\ a_{3y} \\ a_{1z}\ a_{2z}\ a_{3z} \end{pmatrix}, \tag{5.2.3}$$

the volume v of the primitive cell – i.e., the triple scalar product of the vectors – is given by the determinant

$$v = \boldsymbol{a}_1 \cdot (\boldsymbol{a}_2 \times \boldsymbol{a}_3) = [\boldsymbol{a}_1\, \boldsymbol{a}_2\, \boldsymbol{a}_3] = \det A\,. \tag{5.2.4}$$

In terms of the column vectors, the position vector $\boldsymbol{R}_n = n_1\boldsymbol{a}_1 + n_2\boldsymbol{a}_2 + n_3\boldsymbol{a}_3$ of an arbitrary lattice point of the Bravais lattice is

$$\boldsymbol{R}_n = n_1 \begin{pmatrix} a_{1x} \\ a_{1y} \\ a_{1z} \end{pmatrix} + n_2 \begin{pmatrix} a_{2x} \\ a_{2y} \\ a_{2z} \end{pmatrix} + n_3 \begin{pmatrix} a_{3x} \\ a_{3y} \\ a_{3z} \end{pmatrix}. \tag{5.2.5}$$

Using the matrix A, this can be simply rewritten as

$$\boldsymbol{R}_n = \begin{pmatrix} a_{1x}\ a_{2x}\ a_{3x} \\ a_{1y}\ a_{2y}\ a_{3y} \\ a_{1z}\ a_{2z}\ a_{3z} \end{pmatrix} \cdot \begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix} = A \cdot \begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix}. \tag{5.2.6}$$

The matrix B is then defined through the relation

$$BA = 2\pi I\,, \tag{5.2.7}$$

where I is the unit matrix. Denoting the elements of B by

$$B = \begin{pmatrix} b_{1x}\ b_{1y}\ b_{1z} \\ b_{2x}\ b_{2y}\ b_{2z} \\ b_{3x}\ b_{3y}\ b_{3z} \end{pmatrix}, \tag{5.2.8}$$

the vectors

$$\boldsymbol{b}_i = b_{ix}\hat{\boldsymbol{x}} + b_{iy}\hat{\boldsymbol{y}} + b_{iz}\hat{\boldsymbol{z}}\,, \qquad i = 1, 2, 3 \tag{5.2.9}$$

are introduced, which will be written as row vectors

$$\boldsymbol{b}_i = (b_{ix}\ b_{iy}\ b_{iz})\,, \qquad i = 1, 2, 3\,. \tag{5.2.10}$$

Writing the defining equation (5.2.7) in component form, we have

$$b_{ix}a_{jx} + b_{iy}a_{jy} + b_{iz}a_{jz} = 2\pi\delta_{ij}\,. \tag{5.2.11}$$

This can be expressed equivalently in terms of the scalar product of vectors $\boldsymbol{b}_i$ and $\boldsymbol{a}_j$:

$$\boldsymbol{b}_i \cdot \boldsymbol{a}_j = 2\pi\delta_{ij} \,. \tag{5.2.12}$$

Each of the three new vectors is thus perpendicular to the plane spanned by the two *primitive* vectors whose index is different. The vector triplet satisfying this condition is easily found:

$$\boldsymbol{b}_1 = 2\pi\frac{\boldsymbol{a}_2 \times \boldsymbol{a}_3}{\boldsymbol{a}_1 \cdot (\boldsymbol{a}_2 \times \boldsymbol{a}_3)} \,, \quad \boldsymbol{b}_2 = 2\pi\frac{\boldsymbol{a}_3 \times \boldsymbol{a}_1}{\boldsymbol{a}_1 \cdot (\boldsymbol{a}_2 \times \boldsymbol{a}_3)} \,, \quad \boldsymbol{b}_3 = 2\pi\frac{\boldsymbol{a}_1 \times \boldsymbol{a}_2}{\boldsymbol{a}_1 \cdot (\boldsymbol{a}_2 \times \boldsymbol{a}_3)} \,. \tag{5.2.13}$$

Considered as primitive vectors, the $\boldsymbol{b}_i$ span a lattice whose points are given by

$$\boldsymbol{G} = h_1\boldsymbol{b}_1 + h_2\boldsymbol{b}_2 + h_3\boldsymbol{b}_3 \,, \tag{5.2.14}$$

where h_1, h_2, and h_3 are arbitrary integers. In analogy to (5.2.6), these can be written in the form

$$\boldsymbol{G} = (h_1\ h_2\ h_3) \cdot B \,. \tag{5.2.15}$$

The new lattice is called the *reciprocal lattice* of the original lattice, and $\boldsymbol{b}_1$, $\boldsymbol{b}_2$, and $\boldsymbol{b}_3$ are the *primitive vectors of the reciprocal lattice*. In connection with its reciprocal the original lattice is often called the *direct lattice*.

5.2.2 Properties of the Reciprocal Lattice

Transposition of both sides in (5.2.7) gives

$$A^{\mathrm{T}} B^{\mathrm{T}} = 2\pi I \,. \tag{5.2.16}$$

This means that when a lattice is generated using as primitive vectors the $\boldsymbol{b}_i$ – i.e., the columns of the matrix B^{T} –, then the primitive vectors of its reciprocal lattice are given by the rows of the matrix A^{T}, and also

$$\boldsymbol{a}_1 = 2\pi\frac{\boldsymbol{b}_2 \times \boldsymbol{b}_3}{\boldsymbol{b}_1 \cdot (\boldsymbol{b}_2 \times \boldsymbol{b}_3)} \,, \quad \boldsymbol{a}_2 = 2\pi\frac{\boldsymbol{b}_3 \times \boldsymbol{b}_1}{\boldsymbol{b}_1 \cdot (\boldsymbol{b}_2 \times \boldsymbol{b}_3)} \,, \quad \boldsymbol{a}_3 = 2\pi\frac{\boldsymbol{b}_1 \times \boldsymbol{b}_2}{\boldsymbol{b}_1 \cdot (\boldsymbol{b}_2 \times \boldsymbol{b}_3)} \,. \tag{5.2.17}$$

These relations could have also been derived using (5.2.13) for the vectors $\boldsymbol{b}_i$ and making use of the identity (3.3.6). The reciprocal of the reciprocal lattice is thus the original direct lattice.

The primitive cell of the reciprocal lattice is spanned by the primitive vectors $\boldsymbol{b}_i$. Its volume $v_{\mathrm{r}} = \boldsymbol{b}_1 \cdot (\boldsymbol{b}_2 \times \boldsymbol{b}_3)$ is related to the volume v of the direct-lattice primitive cell – given in (5.2.4) – in a particularly simple manner. Expressing the primitive vectors of the reciprocal lattice in terms of their direct-lattice counterparts and using of the vector identity (3.3.6) gives

$$v_{\mathrm{r}} = \boldsymbol{b}_1 \cdot (\boldsymbol{b}_2 \times \boldsymbol{b}_3) = \frac{(2\pi)^3}{\boldsymbol{a}_1 \cdot (\boldsymbol{a}_2 \times \boldsymbol{a}_3)} = \frac{(2\pi)^3}{v} \,. \tag{5.2.18}$$

Alternatively, the definition of the vectors $\boldsymbol{b}_i$ implies that the volume v_r is just the determinant of the matrix B. Making use of (5.2.7) and (5.2.4),

$$v_\mathrm{r} = \det B = \frac{(2\pi)^3}{\det A} = \frac{(2\pi)^3}{v} \,. \tag{5.2.19}$$

Instead of the parallelepiped spanned by the $\boldsymbol{b}_i$, cells of different shape but equal volume may be used in the reciprocal lattice, too. In particular, one might choose Dirichlet's construction, which gave the Wigner–Seitz cell in the direct lattice. In the reciprocal lattice it leads to the Brillouin zone,[8] which will be presented in detail in Chapters 7 and 17.

Expression (5.2.12) for the scalar product of $\boldsymbol{a}_i$ and $\boldsymbol{b}_j$ implies that the product of any reciprocal-lattice vector $\boldsymbol{G}_h = h_1\boldsymbol{b}_1 + h_2\boldsymbol{b}_2 + h_3\boldsymbol{b}_3$ with any direct-lattice vector $\boldsymbol{R}_n = n_1\boldsymbol{a}_1 + n_2\boldsymbol{a}_2 + n_3\boldsymbol{a}_3$ is

$$\boldsymbol{G}_h \cdot \boldsymbol{R}_n = 2\pi(n_1h_1 + n_2h_2 + n_3h_3) \,. \tag{5.2.20}$$

Since the n_i and the h_j are integers, this implies a relation that will be frequently used in the forthcoming:

$$\mathrm{e}^{\mathrm{i}\boldsymbol{G}_h\cdot\boldsymbol{R}_n} = 1 \,. \tag{5.2.21}$$

This relation has an important consequence that will be proved in Appendix C. When a function F is *lattice periodic*, i.e., satisfies the condition

$$F(\boldsymbol{r}) = F(\boldsymbol{r} + \boldsymbol{R}_n) \tag{5.2.22}$$

for any vector $\boldsymbol{R}_n$ of the lattice then its Fourier series contains only the reciprocal-lattice vectors: the sum in

$$F(\boldsymbol{r}) = \sum_{\boldsymbol{G}} F_{\boldsymbol{G}}\, \mathrm{e}^{\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{r}} \tag{5.2.23}$$

is over reciprocal-lattice vectors only.

Finally, another relation between reciprocal-lattice vectors and the direct lattice is worth noting. The definition of the direct-lattice plane with Miller indices (hkl) on page 119 was fairly complicated because the primitive vectors generally form an oblique coordinate system, and the plane (hkl) is not necessarily perpendicular to the direction $[hkl]$ of the vector $h\,\boldsymbol{a}_1 + k\,\boldsymbol{a}_2 + l\,\boldsymbol{a}_3$. On the other hand, the reciprocal-lattice vector

$$\boldsymbol{G}_{hkl} = h\,\boldsymbol{b}_1 + k\,\boldsymbol{b}_2 + l\,\boldsymbol{b}_3 \tag{5.2.24}$$

is readily shown to be directed along the normal to the planes with Miller indices (hkl). It should be recalled that these planes intercept the three crystallographic axes in $m\boldsymbol{a}_1/h$, $m\boldsymbol{a}_2/k$, and $m\boldsymbol{a}_3/l$, where h, k, and l are relatively

[8] L. Brillouin, 1930.

prime, and m is also an integer. This means that the vector $m\boldsymbol{a}_1/h - m\boldsymbol{a}_2/k$ lies in the plane (hkl). The relation between direct- and reciprocal-lattice vectors immediately implies that

$$(h\,\boldsymbol{b}_1 + k\,\boldsymbol{b}_2 + l\,\boldsymbol{b}_3) \cdot \Big(\frac{m}{h}\boldsymbol{a}_1 - \frac{m}{k}\boldsymbol{a}_2\Big) = 0\,, \tag{5.2.25}$$

that is, the two vectors are perpendicular. By the same token, the vector $h\,\boldsymbol{b}_1 + k\,\boldsymbol{b}_2 + l\,\boldsymbol{b}_3$ is perpendicular to other vectors of the plane, e.g., $m\boldsymbol{a}_1/h - m\boldsymbol{a}_3/l$, confirming that $\boldsymbol{G}_{hkl}$ is indeed the normal of the plane.

Making use of this property, the separation d_{hkl} of the neighboring members of the family of planes with Miller indices (hkl) is easily expressed in terms of the magnitude of the reciprocal-lattice vector $\boldsymbol{G}_{hkl}$. Two adjacent planes of the family intercept the $\boldsymbol{a}_1$-axis of the direct lattice in ma_1/h and $(m+1)a_1/h$. The separation of these planes is given by the projection of the vector joining the two intercepts on the normal of the planes:

$$d_{hkl} = \frac{1}{h}\boldsymbol{a}_1 \cdot \frac{\boldsymbol{G}_{hkl}}{|\boldsymbol{G}_{hkl}|} = \frac{2\pi}{|\boldsymbol{G}_{hkl}|}\,. \tag{5.2.26}$$

5.3 Rotations and Reflections

Besides translational symmetry, crystals have further characteristic symmetries. Such symmetry operations are geometrical transformations that leave distances and angles unaffected. Therefore, before turning to the symmetries of Bravais lattices and crystals, we shall present the groups of possible rotations and reflections.

5.3.1 Symmetry Operations and Symmetry Elements

The congruence transformations (isometries) of interest to us are *rotations* through a finite angle α around an axis along the direction of the unit vector $\boldsymbol{n}$, *reflections* in a mirror plane of normal $\boldsymbol{n}$, and *inversion*. Reflection across a line (mirror line) leads to the same result as rotation around the same line through 180°, which is why we shall not discuss it any further. Inversion (i.e., reflection through a point) is, however, treated separately, even though rotation through 180° around an axis $\boldsymbol{n}$ followed or preceded by reflection across a plane of normal $\boldsymbol{n}$ leads to the same result as inversion in the intersection point of the rotation axis and the mirror plane.

The symmetry elements of an object are rotation axes, mirror planes, and/or an inversion center such that the figure is taken into itself by rotations through suitably chosen angles around them or by reflections in them.[9] When

[9] In the spirit of the foregoing, the mirror line is omitted from the previous listing. However, to keep the wording simple, we shall sometimes refer to mirror lines – by which we shall invariably mean rotation axes with 180° operations.

the object is brought into coincidence with itself by a rotation through $2\pi/n$ around an axis, the axis is called an *n-fold symmetry axis*.[10] Two notational conventions are used in the literature for rotations as symmetry operations and rotation axes as symmetry elements. Crystallographic texts dealing with the structural study of crystals almost invariably use the *international* or *Hermann–Mauguin symbols*,[11] while in solid-state physics texts the traditional *Schoenflies symbols*[12] are more common.

In this book we shall use either or both of them. In the present chapter on structure preference will be given to the international notation: this will appear first, followed by the Schoenflies symbol in parentheses. In later chapters dealing with physical properties we shall mostly use the Schoenflies notation.

In the international notation the integer n stands for an n-fold rotation axis, as well as the clockwise rotation through $2\pi/n$ around this axis. For the same concepts Schoenflies proposed the symbol C_n, making reference to the word "cyclic". In other cases, too, even when not marked explicitly, the same notation is used for the symmetry element and the symmetry operation. We shall see later that for crystals n can take on the values 1, 2, 3, 4, and 6 only – however, for the time being we do not impose this restriction. The special case $n = 1$ corresponds to the unit transformation E.[13] For rotational elements the direction of the rotation axis is customarily given by a letter or number label. For example, the notation 2_x (C_{2x}) means that x is a twofold rotation axis, but at the same time it also refers to a 180° rotation around the x-axis. 2_p (C_{2p}), with $p = a, b, c, d, e, f$ stands for 180° rotations around the face diagonals in a cube, while 3_j (C_{3j}), with $j = 1, 2, 3, 4$ stands for 120° rotations around the space diagonals.

The restriction that only rotations through $2\pi/n$ (where n is an integer) and their repetitions should be considered is a consequence of the group property of symmetries. Suppose that the angle of the smallest rotation that takes the system into itself is written as $2\pi/n$. Then k subsequent rotations correspond to a single rotation through $2\pi k/n$ (denoted by n^k or C_n^k). Since rotation through 2π (which is equivalent to the identity transformation) has to be reached in a finite number of steps, n obviously has to be an integer, and only the values $k = 1, 2, \ldots, (n-1)$ correspond to different rotations.

The international notation for *reflections* as symmetry operations and *mirror planes* as symmetry elements is m, while their Schoenflies symbol is C_s or σ. The normal to the mirror plane may appear in the subscript, e.g., m_x (σ_x) refers to reflection in the plane perpendicular to the x-axis. When a mirror plane and a rotation axis are present simultaneously then the orientation of the plane with respect to the axis is also indicated. The notation for a mirror plane perpendicular to an n-fold rotation axis is n/m (σ_h). For mirror planes

[10] $n = 2, 3, 4, 6$-fold axes are also called diad, triad, tetrad, and hexad axes.
[11] C. Hermann, 1928; C.-V. Mauguin, 1931.
[12] A. M. Schoenflies, 1891.
[13] From the German word for unity, *Einheit*.

through the rotation axis Schoenflies used the symbols σ_v and σ_d, where v refers to vertical and d to diagonal. The latter is used for mirror planes that contain a principal rotation axis and bisect the angle between the twofold axes perpendicular to the principal axis.

Inversion as a symmetry operation (which takes point $\boldsymbol{r}$ into $-\boldsymbol{r}$) and the inversion center as a symmetry element are both denoted by $\bar{1}$ (C_i or I). Note that Bravais lattices are always symmetric under inversion since their definition implies that if $\boldsymbol{R}_n$ is a lattice point then so is $-\boldsymbol{R}_n$. We shall make use of this fact when investigating the possible symmetry groups of Bravais lattices.

It has been mentioned that inversion can be considered as a combined symmetry operation. Denoting rotation through angle α around the axis $\boldsymbol{n}$ by $C_{\boldsymbol{n}}(\alpha)$ and reflection in the plane of normal $\boldsymbol{n}$ by $\sigma_{\boldsymbol{n}}$,

$$I = \sigma_{\boldsymbol{n}} C_{\boldsymbol{n}}(\pi) = C_{\boldsymbol{n}}(\pi)\sigma_{\boldsymbol{n}} \,. \tag{5.3.1}$$

It is possible that a system is invariant neither under 180° rotation nor under reflection – only under the combination of the two. A similar situation may arise with other operations as well. Some objects will be taken into themselves only when rotation through $2\pi/n$ is followed by an inversion in a point on the rotation axis or by a reflection in a mirror plane perpendicular to the axis. These improper rotations are symmetry operations of the second kind; the first one is called *rotation–inversion* or *rotoinversion*, and the second – *rotation–reflection* or *rotoreflection*. The order of the two operations is immaterial. Thus rotoreflection of an object is the rotation of its mirror image. The corresponding symmetry elements are the *rotation–inversion axis* and the *rotation–reflection axis*.[14]

When listing the symmetry operations of a given system, either rotation–inversion or rotation–reflection has to be included – but not both, as they are related to each other in a particularly simple way. Rotation–inversion and rotation–reflection lead to the same result when the two angles of rotation differ by π.

Rotation–inversion – that is, rotation through $2\pi/n$ around a rotation axis, followed by an inversion in a point of the rotation axis – is considered as the fundamental operation in the international crystallographic system. Both the operation and the corresponding symmetry element, the rotation–inversion axis are denoted by $\bar{n}$ (C_{ni}). Since the order of rotation and inversion can be interchanged, $\bar{n} = n\,\bar{1} = \bar{1}\,n$ ($C_{ni} = C_n\,I = I\,C_n$). This symmetry may appear even when neither rotation n (C_n) nor inversion is a symmetry of the system alone. However, this is not the case for $n = 1$ and $n = 3$. $n = 1$ corresponds to a pure inversion, while for $n = 3$ $\bar{3}^2 = 3^{-1}$ ($C_{3i}^2 = C_3^{-1}$), $\bar{3}^3 = \bar{1}$ ($C_{3i}^3 = I$), and $\bar{3}^4 = 3$ ($C_{3i}^4 = C_3$). The twofold operation $\bar{2}$ (C_{2i}) is not a new one, either,

[14] 2-, 3-, 4-, and 6-fold rotation–inversion axes are also called *inversion diad, triad, tetrad, and hexad axes*, respectively.

since it corresponds to a reflection in the plane perpendicular to the rotation axis: $\bar{2} \equiv m$ ($C_{2i} \equiv C_s$).

Following Schoenflies' proposition, rotation–reflection – that is, rotation through $2\pi/n$ followed by a reflection in a plane perpendicular to the axis of rotation – is often considered to be a fundamental operation. This symmetry operation and the corresponding symmetry element, the rotoreflection axis are both denoted by S_n,[15] or $\widetilde{n}$ in the international notation. Because of the interchangeability of the two transformations $S_n = \sigma_h C_n = C_n \sigma_h$ ($\widetilde{n} = n\,m = m\,n$). Specifically, S_1 ($\widetilde{1}$) is a reflection and S_2 ($\widetilde{2}$) is an inversion.

As mentioned above, the two symmetry operations of the second kind are not independent of one another, e.g., (i) $\widetilde{1} = \bar{2} \equiv m$ ($S_1 = C_{2i} \equiv \sigma$); (ii) $\widetilde{2} = \bar{1}$ ($S_2 = C_i \equiv I$); (iii) $\widetilde{3} = \bar{6}^{-1}$ ($S_3 = C_{6i}^{-1}$); (iv) $\widetilde{4} = \bar{4}^{-1}$ ($S_4 = C_{4i}^{-1}$); (iv) $\widetilde{6} = \bar{3}^{-1}$ ($S_6 = C_{3i}^{-1}$).

5.3.2 Point Groups

With the sole exception of spheres, all figures are invariant only under a finite of infinite subgroup of the above symmetry operations. Possible symmetry operations constitute a group in the algebraic sense: when each of two symmetry operations takes a figure into itself then the succession of the two (which can be defined as the group multiplication operation) is also a symmetry. These elements make up the symmetry group of the figure. Before identifying the group of possible rotation and reflection symmetries of crystal lattices, consider finite geometrical figures that have rotation axes, mirror planes, and/or inversion centers. Due to finiteness and hence the lack of translational symmetry, one point of the figure always remains invariant. Rotation axes and mirror planes must all go through this point, and if the figure has an inversion center then it must be in the same point as well. The group of symmetry transformations that leave this point invariant is called the *point group* of the figure. For d-dimensional figures these are collectively denoted by G_0^d.

Point groups may also be defined as the discrete subgroups of the full orthogonal group $\mathrm{O}(3)$[16] (or, in the d-dimensional case, of the group $\mathrm{O}(d)$): they consist of rotations through $2\pi/n$, reflections, inversions, rotation–inversions, and their combinations.

Two-Dimensional Point Groups (G_0^2)

The set of possible symmetry operations for two-dimensional (plane) figures consists of rotations around the axis perpendicular to the plane, reflection in

[15] S stands for "Spiegel", the German word for mirror.

[16] The full rotation group (or special orthogonal group) $\mathrm{SO}(3)$ consists of rotations that leave a specific point of space – the fixed point – invariant. In addition to these, the full orthogonal group contains the inversion operation and rotation–inversions.

a line within the plane, and inversion in a point of the plane. The last one can be ignored as it corresponds to a rotation through 180°. Following the international notation, we shall denote n-fold rotations by n, and mirror lines by m.

The simplest point groups consist of rotations through $2\pi/n$ and its multiples. These n-element groups are also denoted by n. For the symmetries of finite figures all integer values n are allowed.

The two-element point group consisting of a reflection in a line and the square of this operation (i.e., the identity element) is denoted by m.

Now consider those groups that contain mirror lines and a rotation axis that is perpendicular to the plane. Then any line obtained from a mirror line via rotation through $2\pi/n$ around the n-fold axis is also a mirror line. For n odd, the angular separation of the n mirror lines obtained this way is just π/n. The group thus has $2n$ elements and is denoted by nm.

For n even, rotations of a mirror line will yield only $n/2$ different mirror lines whose angular separation is $2\pi/n$. Nevertheless in this case there must exist another set of mirror lines – the angle bisectors of the previously obtained lines. This is so because the composition of a reflection in a mirror line and a rotation through $2\pi/n$ is equivalent to a reflection in a mirror line that makes an angle π/n with the original mirror line. Thus there are two independent sets of mirror lines. This is expressed by the notation nmm of such groups.

Three-Dimensional Point Groups (G_0^3)

We shall now turn to the much more interesting case of three dimensions where point groups can be divided into two large classes. Point groups of the first kind consist only of rotations, while those of the second kind contain reflections as well.[17] This means that point groups of the first kind are discrete subgroups of the rotation group SO(3).

Again, two equally common notational conventions are used in the literature. In the international notation numbers – more specifically: numbers giving the order of the rotation axes – refer to rotational symmetry, and the letter m to (mirror) reflection. For example, 432 means that the point group contains four-, three-, and twofold rotation axes, while $4/m$ means that the plane perpendicular to the fourfold rotation axis is a mirror plane. Point groups have a full and a short notation. For the symmetry group of a cube – where planes perpendicular to the fourfold axes parallel to the edges and to the twofold axes parallel to the face diagonals are mirror planes, and space diagonals are threefold rotation axes as well as threefold rotation–inversion axes –, the point

[17] This distinction between point groups of the first and second kinds is justified by the fact that rotation conserves the "handedness" (chirality) of the object, while reflection, inversion, rotoreflection and rotoinversion change it. The latter are called enantiomorphous operations – from the Greek word, $\epsilon\nu\alpha\nu\tau\iota o\varsigma$ (enantios) for opposite.

group can be specified as $4/m\,\overline{3}\,2/m$. However, it is customary to use the short notation $m\overline{3}m$. In what follows, we shall give the short version only.

The Schoenflies notation uses the symbols of symmetry operations (C, S) or letters that refer to figures with the appropriate symmetry (D = dihedral, T = tetrahedral, O = octahedral, I or Y = icosahedral), and optionally an index (h = horizontal, v = vertical, d = diagonal). Below we shall first give the short international notation and then the Schoenflies symbol in parentheses.

Point Groups of the First Kind

Several intersecting rotation axes are possible in point groups of the first kind. Symmetries of one axis determine the number as well as relative spatial orientation of the other axes. Point groups of the first kind exist in five types. These correspond to the rotational symmetry groups of the simple geometric figures shown in Fig. 5.12. These figures may possess further symmetries, e.g. mirror planes which will be ignored here as we are concerned with rotations only.

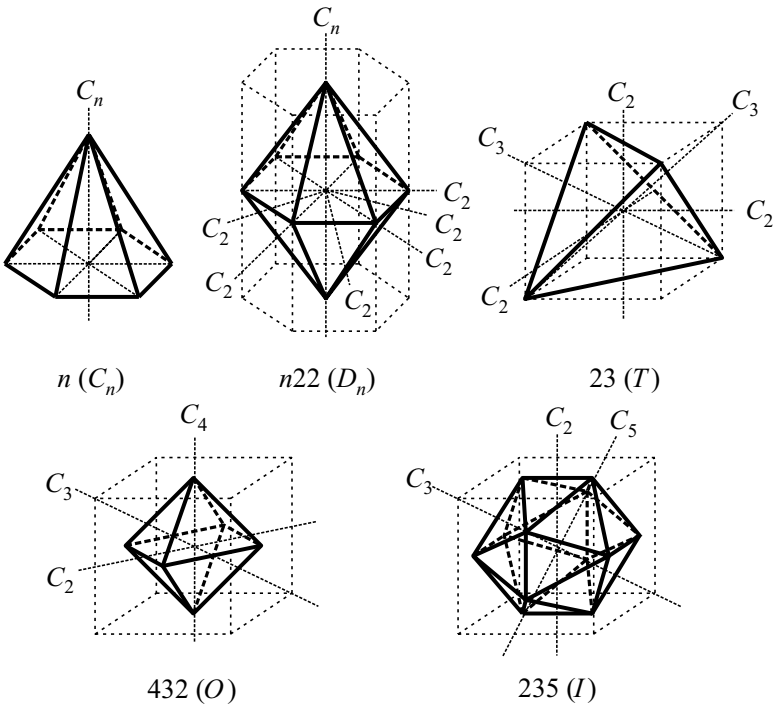


Fig. 5.12. Finite figures showing the symmetries of the five point groups of the first kind

1. *Cyclic or uniaxial group*: n (C_n). The group of rotations through angles $2\pi k/n$ around a rotation axis. It has n elements.
2. *Dihedral group*: $n22$ for n even and $n2$ for n odd. (The Schoenflies symbol is D_n in both cases.) A group with $2n$ elements that contains rotations around an n-fold rotation axis as well as rotations about n twofold axes that are perpendicular to the first axis. The difference between n odd and n even – which lies at the root of the difference in notation – is whether the twofold axes are obtained from the same axis or two separate axes by n-fold rotation. In the $n = 2$ case the notations D_2 and V are both common.
3. *Tetrahedral group*: 23 (T). A group with 12 elements that contains rotations around four threefold and three twofold axes that all pass through a common point. The relative orientations of the rotation axes are such that they transform a regular tetrahedron into itself.
4. *Octahedral group*: 432 (O). A group with 24 elements, which contains all those two-, three- and fourfold rotations that take a cube or an octahedron inscribed in a cube into itself. These symmetry operations are the same as the rotations that take a cube into itself. In what follows, we shall frequently use these symmetries, therefore we shall also list them separately in Table 5.1.

Besides the Schoenflies symbols, the Jones symbol – also called the *faithful representation* – is also given in the table. Using the latter, the symmetry operation is characterized by the three relations

$$\begin{aligned} x' &= f(x,y,z)\,, \\ y' &= g(x,y,z)\,, \\ z' &= h(x,y,z) \end{aligned} \tag{5.3.2}$$

that specify the point $\boldsymbol{r}' = x'\hat{\boldsymbol{x}} + y'\hat{\boldsymbol{y}} + z'\hat{\boldsymbol{z}}$ into which $\boldsymbol{r} = x\hat{\boldsymbol{x}} + y\hat{\boldsymbol{y}} + z\hat{\boldsymbol{z}}$ is taken by the symmetry operation. It is customary to have $\bar{x} = -x$. It is particularly simple to find the product of two group elements – which is obtained via the composition of the two symmetry operations – in the Jones faithful representation. For example, by making use of $C_{34}^2 = (\bar{y}, \bar{z}, x)$ and $C_{2c} = (z, \bar{y}, x)$ one is lead to

$$C_{2c}C_{34}^2(x\hat{\boldsymbol{x}} + y\hat{\boldsymbol{y}} + z\hat{\boldsymbol{z}}) = C_{2c}(\bar{y}\hat{\boldsymbol{x}} + \bar{z}\hat{\boldsymbol{y}} + x\hat{\boldsymbol{z}}) = x\hat{\boldsymbol{x}} + z\hat{\boldsymbol{y}} + \bar{y}\hat{\boldsymbol{z}}\,,$$

that is, the Jones notation for $C_{2c}C_{34}^2$ is $x, z, \bar{y}$ – according to Table 5.1 this is the faithful representation of C_{4x}^3 –, thus $C_{2c}C_{34}^2 = C_{4x}^3$, while

$$C_{34}^2C_{2c}(x\hat{\boldsymbol{x}} + y\hat{\boldsymbol{y}} + z\hat{\boldsymbol{z}}) = C_{34}^2(z\hat{\boldsymbol{x}} + \bar{y}\hat{\boldsymbol{y}} + x\hat{\boldsymbol{z}}) = y\hat{\boldsymbol{x}} + \bar{x}\hat{\boldsymbol{y}} + z\hat{\boldsymbol{z}}\,,$$

that is, $C_{34}^2C_{2c}$ is represented by $y, \bar{x}, z$, which corresponds to C_{4z}^3, and so $C_{34}^2C_{2c} = C_{4z}^3$.

Table 5.1. Rotations around axes through the cube center that take the cube into itself

Schoenflies symbol	Jones symbol	Description of the symmetry operation
E	x, y, z	identity element
C_{2x}	$x, \bar{y}, \bar{z}$	rotation through 180° around [100]
C_{2y}	$\bar{x}, y, \bar{z}$	rotation through 180° around [010]
C_{2z}	$\bar{x}, \bar{y}, z$	rotation through 180° around [001]
C_{31}	z, x, y	rotation through 120° around [111]
C_{32}	$\bar{z}, x, \bar{y}$	rotation through 120° around $[\bar{1}\bar{1}1]$
C_{33}	$\bar{z}, \bar{x}, y$	rotation through 120° around $[1\bar{1}\bar{1}]$
C_{34}	$z, \bar{x}, \bar{y}$	rotation through 120° around $[\bar{1}1\bar{1}]$
C_{31}^2	y, z, x	rotation through 240° around [111]
C_{32}^2	$y, \bar{z}, \bar{x}$	rotation through 240° around $[\bar{1}\bar{1}1]$
C_{33}^2	$\bar{y}, z, \bar{x}$	rotation through 240° around $[1\bar{1}\bar{1}]$
C_{34}^2	$\bar{y}, \bar{z}, x$	rotation through 240° around $[\bar{1}1\bar{1}]$
C_{4x}	$x, \bar{z}, y$	rotation through 90° around [100]
C_{4y}	$z, y, \bar{x}$	rotation through 90° around [010]
C_{4z}	$\bar{y}, x, z$	rotation through 90° around [001]
C_{4x}^3	$x, z, \bar{y}$	rotation through 270° around [100]
C_{4y}^3	$\bar{z}, y, x$	rotation through 270° around [010]
C_{4z}^3	$y, \bar{x}, z$	rotation through 270° around [001]
C_{2a}	$y, x, \bar{z}$	rotation through 180° around [110]
C_{2b}	$\bar{y}, \bar{x}, \bar{z}$	rotation through 180° around $[1\bar{1}0]$
C_{2c}	$z, \bar{y}, x$	rotation through 180° around [101]
C_{2d}	$\bar{x}, z, y$	rotation through 180° around [011]
C_{2e}	$\bar{z}, \bar{y}, \bar{x}$	rotation through 180° around $[10\bar{1}]$
C_{2f}	$\bar{x}, \bar{z}, \bar{y}$	rotation through 180° around $[01\bar{1}]$

5. *Icosahedral group*: 235 (I or Y). A group with 60 elements, which contains 6 fivefold, 10 threefold and 15 twofold axes arranged in such a way that they take a regular pentagonal dodecahedron (a polyhedron bounded by 12 congruent regular pentagonal faces) or a regular icosahedron (a polyhedron bounded by 20 congruent regular triangular faces) into itself.

Point Groups of the Second Kind

Besides rotations, point groups of the second kind also contain inversions, reflections, and rotation–inversions or rotation–reflections. Some of these groups are obtained by adding inversion and its products with rotations to point groups of the first kind. In other cases rotations are complemented by reflections. Comparison of the international (Hermann–Mauguin) and Schoen-

flies notations is cumbersome because the fundamental operation is chosen as rotation–inversion in the former and rotation–reflection in the latter. That is why the point group $\bar{n}$, which contains rotation–inversions, appears as S_{2n} in the Schoenflies system when n is odd, as S_n when n is divisible by 4, and as $C_{2l+1,h}$ when $n = 2(2l + 1)$. For notational simplicity, we shall exceptionally give the Schoenflies notation first. Some simple relations between the two notations are given in Table 5.2.

1. S_{2n} ($\overline{2n}$ if n is even and $\bar{n}$ if n is odd): a group with $2n$ elements, which contains rotation–reflections around an n-fold axis. This group contains the point group C_n as its subgroup. For $n = 1$ the group S_2 has two elements, the identity element E and inversion I. The notation C_i is also used for this group.
2. C_{nh} (n/m if n is even and $\overline{2n}$ if n is odd): a group with $2n$ elements, which contains rotations around an n-fold axis and a mirror plane that is perpendicular to it. For $n = 1$ the group C_{1h} has two elements, the identity element E and the reflection σ_h (m). The notations C_s and S_1 (m) are also used for this group.
3. C_{nv} (nmm if n is even and nm if n is odd): a group with $2n$ elements, which contains rotations around an n-fold axis and n mirror planes, each of which passes through the axis. This is the group of symmetries for a right pyramid whose base is a regular n-gon.
4. D_{nh} (n/mmm if n is even and ($\overline{2n}$)$2m$ if n is odd): a group with $4n$ elements, which contains the symmetry elements of the point group D_n plus a mirror plane perpendicular to the n-fold axis. This is the group of symmetries for a right prism whose base is a regular n-gon. Of particular interest is the group D_{6h}, which contains the 24 symmetries of the regular

Table 5.2. Connections between international and Schoenflies symbols for point groups

	$n = 2l + 1$	$n = 4l + 2$	$n = 4l$
n	C_n	C_n	C_n
$\bar{n}$	S_{2n}	$C_{\frac{1}{2}nh}$	S_n
n/m		C_{nh}	C_{nh}
$n2$	D_n		
$n22$		D_n	D_n
nm	C_{nv}		
nmm		C_{nv}	C_{nv}
$\bar{n}m$	D_{nd}		
$\bar{n}2m$		$D_{\frac{1}{2}nd}$	$D_{\frac{1}{2}nd}$
n/mmm		D_{nh}	D_{nh}

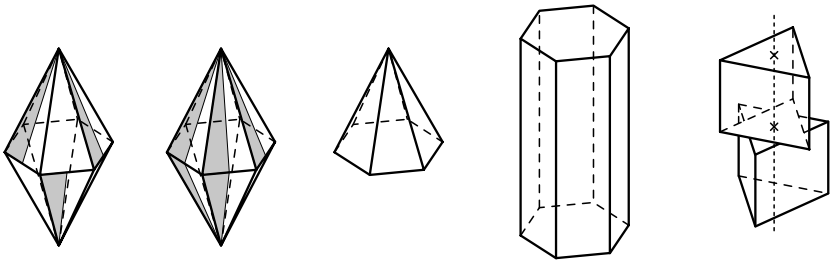
hexagonal prism. The 24 operations are listed in Table 5.3. In the Jones notation the components are given in a coordinate system that is particularly well suited to hexagonal symmetry: the x- and y-axes make an angle of 120°. In the $n = 2$ case the notations D_{2h} and V_h are also used.

5. D_{nd} ($(\overline{2n})2m$ if n is even and $\bar{n}m$ if n is odd): a group with $4n$ elements, which contains the symmetry elements of the point group D_n plus mirror planes containing the n-fold axis, which bisect the angles between the twofold axes. This is the group of symmetries for a figure in which a regular n-gonal prism is stacked directly above another, and rotated with respect to it. In the $n = 2$ case the notations D_{2d} and V_d are also used.

Figure 5.13 shows finite objects whose full symmetry group is one of the above five point groups.

Table 5.3. Rotation and reflection symmetries of a regular hexagonal prism

Schoenflies symbol	Jones symbol	Description of the symmetry operation
E	x, y, z	identity element
C_6	$x - y, x, z$	rotation through 60° around [001]
C_3	$\bar{y}, x - y, z$	rotation through 120° around [001]
C_2	$\bar{x}, \bar{y}, z$	rotation through 180° around [001]
C_3^2	$-x + y, \bar{x}, z$	rotation through 240° around [001]
C_6^5	$y, -x + y, z$	rotation through 300° around [001]
C_{2a}	$-x + y, y, \bar{z}$	rotation through 180° around [120]
C_{2b}	$x, x - y, \bar{z}$	rotation through 180° around [210]
C_{2c}	$\bar{y}, \bar{x}, \bar{z}$	rotation through 180° around $[1\bar{1}0]$
C_{2d}	$x - y, \bar{y}, \bar{z}$	rotation through 180° around [100]
C_{2e}	$\bar{x}, -x + y, \bar{z}$	rotation through 180° around [010]
C_{2f}	$y, x, \bar{z}$	rotation through 180° around [110]
I	$\bar{x}, \bar{y}, \bar{z}$	inversion
$S_3^2 = I \cdot C_6$	$-x + y, \bar{x}, \bar{z}$	rotoreflection through 240° around [001]
$S_6^5 = I \cdot C_3$	$y, -x + y, \bar{z}$	rotoreflection through 300° around [001]
$\sigma_h = I \cdot C_2$	$x, y, \bar{z}$	reflection in the (001) plane
$S_6 = I \cdot C_3^2$	$x - y, x, \bar{z}$	rotoreflection through 60° around [001]
$S_3 = I \cdot C_6^5$	$\bar{y}, x - y, \bar{z}$	rotoreflection through 120° around [001]
$\sigma_a = I \cdot C_{2a}$	$x - y, \bar{y}, z$	reflection in the (120) plane
$\sigma_b = I \cdot C_{2b}$	$\bar{x}, -x + y, z$	reflection in the (210) plane
$\sigma_c = I \cdot C_{2c}$	y, x, z	reflection in the $(1\bar{1}0)$ plane
$\sigma_d = I \cdot C_{2d}$	$-x + y, y, z$	reflection in the (100) plane
$\sigma_e = I \cdot C_{2e}$	$x, x - y, z$	reflection in the (010) plane
$\sigma_f = I \cdot C_{2f}$	$\bar{y}, \bar{x}, z$	reflection in the (110) plane

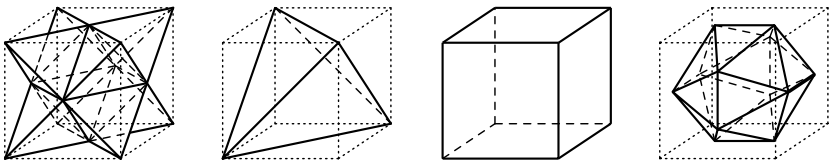


S_6 $(\bar{3})$ C_{6h} $(6/m)$ C_{6v} $(6mm)$ D_{6h} $(6/mm)$ D_{3d} $(\bar{3}m)$

Fig. 5.13. Objects possessing the symmetries of some point groups of the second kind

6. T_h $(m\bar{3})$: a group with 24 elements, which contains the symmetries of the point group T plus inversion. $T_h = T \otimes C_i$ $(m\bar{3} = 23 \otimes \bar{1})$.
7. T_d $(\bar{4}3m)$: a group with 24 elements, which contains all the symmetry operations of a regular tetrahedron. $T_d = T \otimes C_s$ $(\bar{4}3m = 23 \otimes m)$.
8. O_h $(m\bar{3}m)$: a group with 48 elements, which contains all the symmetry operations of a cube. The 24 rotational symmetries of the cube are complemented by 24 new operations. Some of the latter are reflections, others are rotation–reflections. These are listed in Table 5.4. Note that these symmetry operations are obtained from the elements of the point group O (432) listed in Table 5.1 by multiplication by the inversion operation. $O_h = O \otimes C_i$ $(m\bar{3}m = 432 \otimes \bar{1})$.
9. I_h or Y_h $(m\bar{3}\bar{5})$: A group with 120 elements, which contains all the symmetry operations of a regular pentagonal dodecahedron or a regular icosahedron. $I_h = I \otimes C_i$ $(m\bar{3}\bar{5} = 235 \otimes \bar{1})$.

Figure 5.14 shows finite objects whose full symmetry group is one of the above point groups of the second kind, T_h, T_d, O_h, or I_h.



T_h $(m\bar{3})$ T_d $(\bar{4}3m)$ O_h $(m\bar{3}m)$ I_h $(m\bar{3}\bar{5})$

Fig. 5.14. Objects possessing the symmetries of some point groups of the second kind

Table 5.4. Reflection symmetries of a cube

Schoenflies symbol	Jones symbol	Description of the symmetry operation
I	$\bar{x}, \bar{y}, \bar{z}$	inversion
$\sigma_x = I \cdot C_{2x}$	$\bar{x}, y, z$	reflection in the (100) plane
$\sigma_y = I \cdot C_{2y}$	$x, \bar{y}, z$	reflection in the (010) plane
$\sigma_z = I \cdot C_{2z}$	$x, y, \bar{z}$	reflection in the (001) plane
$S_{61}^5 = I \cdot C_{31}$	$\bar{z}, \bar{x}, \bar{y}$	rotoreflection through 300° around [111]
$S_{62}^5 = I \cdot C_{32}$	$z, \bar{x}, y$	rotoreflection through 300° around $[\bar{1}\bar{1}1]$
$S_{63}^5 = I \cdot C_{33}$	$z, x, \bar{y}$	rotoreflection through 300° around $[1\bar{1}\bar{1}]$
$S_{64}^5 = I \cdot C_{34}$	$\bar{z}, x, y$	rotoreflection through 300° around $[\bar{1}1\bar{1}]$
$S_{61} = I \cdot C_{31}^2$	$\bar{y}, \bar{z}, \bar{x}$	rotoreflection through 60° around [111]
$S_{62} = I \cdot C_{32}^2$	$\bar{y}, z, x$	rotoreflection through 60° around $[\bar{1}\bar{1}1]$
$S_{63} = I \cdot C_{33}^2$	$y, \bar{z}, x$	rotoreflection through 60° around $[1\bar{1}\bar{1}]$
$S_{64} = I \cdot C_{34}^2$	$y, z, \bar{x}$	rotoreflection through 60° around $[\bar{1}1\bar{1}]$
$S_{4x}^3 = I \cdot C_{4x}$	$\bar{x}, z, \bar{y}$	rotoreflection through 270° around [100]
$S_{4y}^3 = I \cdot C_{4y}$	$\bar{z}, \bar{y}, x$	rotoreflection through 270° around [010]
$S_{4z}^3 = I \cdot C_{4z}$	$y, \bar{x}, \bar{z}$	rotoreflection through 270° around [001]
$S_{4x} = I \cdot C_{4x}^3$	$\bar{x}, \bar{z}, y$	rotoreflection through 90° around [100]
$S_{4y} = I \cdot C_{4y}^3$	$z, \bar{y}, \bar{x}$	rotoreflection through 90° around [010]
$S_{4z} = I \cdot C_{4z}^3$	$\bar{y}, x, \bar{z}$	rotoreflection through 90° around [001]
$\sigma_a = I \cdot C_{2a}$	$\bar{y}, \bar{x}, z$	reflection in the (110) plane
$\sigma_b = I \cdot C_{2b}$	y, x, z	reflection in the $(\bar{1}10)$ plane
$\sigma_c = I \cdot C_{2c}$	$\bar{z}, y, \bar{x}$	reflection in the (101) plane
$\sigma_d = I \cdot C_{2d}$	$x, \bar{z}, \bar{y}$	reflection in the (011) plane
$\sigma_e = I \cdot C_{2e}$	z, y, x	reflection in the $(\bar{1}01)$ plane
$\sigma_f = I \cdot C_{2f}$	x, z, y	reflection in the $(0\bar{1}1)$ plane

5.4 Rotation and Reflection Symmetries in Crystals

After the presentation of point groups we shall now turn back to crystals and examine the rotations and reflections, or, more precisely, the point groups they might possess. As a first step, we shall study the rotation and reflection symmetries of Bravais lattices. When the Bravais lattice is taken into itself by a rotation or reflection then so is the Wigner–Seitz cell because of its construction. The group of symmetries of a Bravais lattice is thus the group of symmetries of a finite geometrical figure – that is, one of the point groups.

5.4.1 Rotation Symmetries of Bravais Lattices

It has been mentioned that Bravais lattices always possess inversion symmetry (centrosymmetry). Therefore only those point groups need to be considered

that contain inversion.[18] Moreover, the requirement that the lattice should also have discrete translational symmetries imposes a severe restriction on the allowed rotation angles. Indeed, translation symmetry does not allow for invariance under rotations through arbitrary angles $2\pi/n$. We shall demonstrate that if (in addition to the translational symmetry that is due to its periodicity) the lattice possesses rotational symmetry – that is, if there exists an axis such that rotation through an angle φ around it moves the lattice into itself – then the angle of rotation must be $\pm 60°$, $\pm 90°$, $\pm 120°$, or $180°$. We shall first prove this statement for planar lattices, and then show that only two-, three-, four-, and sixfold rotation axes can exist in three-dimensional crystal lattices as well.

Consider a planar (two-dimensional) Bravais lattice. If there exists a rotation axis that is perpendicular to the plane then there exist an infinite number of such axes – whose parallel displacements relative to each other are given by lattice vectors $\boldsymbol{t}_n$. These axes cannot be considered as essentially different, therefore a single one has to be chosen among them. On the other hand, it is conceivable that there are several independent nonparallel rotation axes. Their positions are not arbitrary: they must pass through special points of the primitive cell, e.g., vertices, edge centers, or face centers. More general positions are excluded since the requirements that the lattice should be invariant under translations and rotations could not be fulfilled simultaneously otherwise.

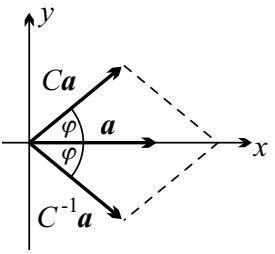


Fig. 5.15. Transformation of a primitive lattice vector under a rotation

Consider a rotation axis that goes through a vertex of the primitive cell, and choose the coordinate system in such a way that the shorter primitive vector, denoted by $\boldsymbol{a}$, should be directed along the x-axis. In this coordinate system $\boldsymbol{a} = (a, 0, 0)$. Now perform the rotations $C(\varphi)$ and $C(-\varphi) = C^{-1}(\varphi)$ – through angles φ and $-\varphi$ –, as shown in Figure 5.15. The resulting vectors are

[18] In two dimensions inversion is equivalent to a rotation through 180° around an axis perpendicular to the plane, thus planar Bravais lattices always possess a twofold rotation axis.

$$\begin{aligned} \boldsymbol{a}' &= C(\varphi)\boldsymbol{a} = a(\cos\varphi, \sin\varphi, 0)\,, \\ \bar{\boldsymbol{a}}' &= C^{-1}(\varphi)\boldsymbol{a} = a(\cos\varphi, -\sin\varphi, 0)\,. \end{aligned} \tag{5.4.1}$$

If rotation C through angle φ is a symmetry of the lattice, then so is C^{-1}, and both points are in the lattice. Moreover, the vector

$$\boldsymbol{a}' + \bar{\boldsymbol{a}}' = a(2\cos\varphi, 0, 0) \tag{5.4.2}$$

must also point to a lattice site. Since this vector is along the direction of $\boldsymbol{a}$, and $\boldsymbol{a}$ was chosen as the shortest lattice vector in the (x, y) plane, the previously obtained vector has to be an integral multiple of $\boldsymbol{a}$. So, for a rotation through angle φ to be a symmetry operation, we must have

$$2\cos\varphi = 2,\ 1,\ 0,\ -1,\ \text{or}\ -2\,. \tag{5.4.3}$$

The corresponding angles are

$$\varphi = 0,\ \ \pm\frac{\pi}{3},\ \ \pm\frac{\pi}{2},\ \ \pm\frac{2\pi}{3},\ \ \pi\,. \tag{5.4.4}$$

Only those angles appeared that correspond to two-, three-, four-, and sixfold rotations. Therefore only those point groups can be the symmetry groups of a Bravais lattice that contain only such rotations. The absence of fivefold rotations can also be explained by the impossibility of tiling the plane perfectly with regular pentagons, as it was mentioned on page 116. While fivefold symmetry cannot be present in periodic structures, it may manifest itself in some patterns that are regular in a broader sense. Such quasiperiodic structures – quasicrystals – will be discussed in Chapter 10.

A similar path is followed in the case of three-dimensional crystals. The origin of the coordinate system is now chosen at a lattice point, and the z-axis is directed along the axis of rotation. It is a straightforward matter to show that the existence of a rotation axis implies that some – not necessarily primitive – lattice vectors lie in the (x, y) plane, which is perpendicular to the rotation axis. Consider the rotation of the primitive vector $\boldsymbol{a}_i$ around the z-axis through an angle φ. Provided this rotation is a symmetry of the lattice, the endpoint of the rotated vector is also at a lattice site. Thus the vector $\boldsymbol{a}_i' - \boldsymbol{a}_i$ is an element of the group of translations. Owing to its construction, it is obviously perpendicular to the rotation axis, that is, it lies in the (x, y) plane. Now consider the shortest lattice vector that starts at the rotation axis and is perpendicular to it. Denoting this by $\boldsymbol{a}$, the procedure seen in the two-dimensional case can be repeated. Finally, a contradiction is found whenever the rotation axis is not a two-, three-, four-, or sixfold axis.

5.4.2 Crystallographic Point Groups

According to the foregoing, crystals can be invariant only under rotations through angles $\pi/3$, $\pi/2$, or an integral multiple of these. This is the reason

why crystallography is concerned only with those point groups that contain two-, three-, four-, or sixfold rotation axes. In the two-dimensional case this implies that only ten point groups need to be studied (1, 2, 3, 4, 6, m, $2mm$, $3m$, $4mm$, and $6mm$). Figure 5.16 shows some simple figures whose symmetry groups are these point groups.

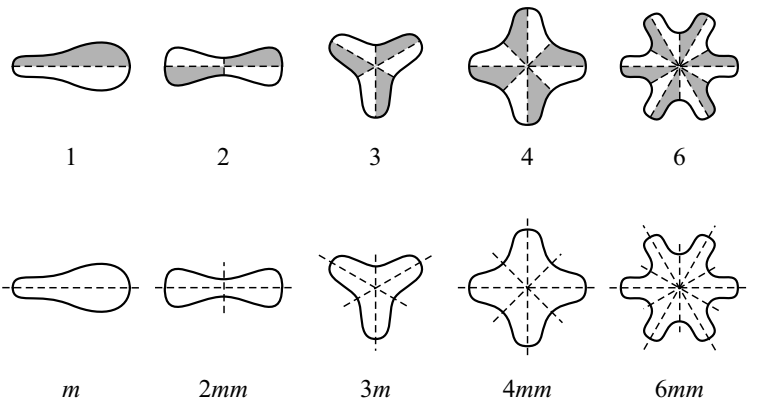


Fig. 5.16. Finite plane figures whose symmetry groups are the ten planar crystallographic point groups

In the three-dimensional case the number of crystallographic point groups – in which only two-, three-, four-, and sixfold rotations are allowed – is 32. These point groups, along with their symmetry operations, are listed in Table 5.5. The reason for dividing point groups into seven classes will transpire later. The relative spatial orientation of the symmetry elements of each group can be deduced from the information provided in the general description. Apart from those of the trigonal and hexagonal systems, all point groups can be obtained by specifying every possible subgroup of the group $m\bar{3}m$ (O_h), which contains all the rotation and reflection symmetries of a cube.[19] Point groups of the trigonal and hexagonal systems can be written down by identifying the subgroups of the point group $6/mmm$ (D_{6h}), which contains the symmetry operations of a regular hexagonal prism.[20]

5.4.3 Crystal Systems and Bravais Groups

In the foregoing we have listed 10 two-dimensional and 32 three-dimensional point groups that might be compatible with the conditions imposed by translational symmetries on the allowed rotational symmetries. However, not each of them occurs as the point group of a Bravais lattice. This is because two more

[19] These rotations and reflections are listed in Tables 5.1 and 5.4.
[20] The 12 rotations and 12 (rotation)–reflections are listed in Table 5.3.

Table 5.5. International and Schoenflies symbols for the 32 crystallographic point groups and their symmetry operations. Subscripts m, j, and p stand for $m = x, y, z$, $j = 1, 2, 3, 4$, and $p = a, b, c, d, e, f$

System	Notation		Symmetry operations
Triclinic	1	C_1	E
	$\bar{1}$	$C_i \equiv S_2$	E, I
Monoclinic	2	C_2	E, C_{2z}
	$\bar{2} \equiv m$	$C_s \equiv C_{1h}$	E, σ_z
	$2/m$	C_{2h}	$C_2 \otimes C_i$
Orthorhombic	222	$D_2 \equiv V$	E, C_{2x}, C_{2y}, C_{2z}
	$mm2$	C_{2v}	E, C_{2z}, σ_x, σ_y
	mmm	$D_{2h} \equiv V_h$	$D_2 \otimes C_i$
Tetragonal	4	C_4	E, C_{4z}, C_{2z}, C_{4z}^3
	$\bar{4}$	S_4	E, S_{4z}, C_{2z}, S_{4z}^3
	$4/m$	C_{4h}	$C_4 \otimes C_i$
	422	D_4	E, C_{4z}, C_{2z}, C_{4z}^3 C_{2x}, C_{2y}, C_{2a}, C_{2b}
	$4mm$	C_{4v}	E, C_{4z}, C_{2z}, C_{4z}^3 σ_x, σ_y, σ_a, σ_b
	$\bar{4}2m$	$D_{2d} \equiv D_d$	E, S_{4z}, C_{2z}, S_{4z}^3 C_{2x}, C_{2y}, σ_a, σ_b
	$4/mmm$	D_{4h}	$D_4 \otimes C_i$
Rhombohedral	3	C_3	E, C_3, C_3^2
(Trigonal)	$\bar{3}$	$C_{3i} \equiv S_6$	$C_3 \otimes C_i$
	32	D_3	E, C_3, C_3^2, C_{2a}, C_{2b}, C_{2c}
	$3m$	C_{3v}	E, C_3, C_3^2, σ_a, σ_b, σ_c
	$\bar{3}m$	D_{3d}	$D_3 \otimes C_i$
Hexagonal	6	C_6	E, C_6, C_3, C_2, C_3^2, C_6^5
	$\bar{6}$	C_{3h}	E, S_3, S_3^2, C_3, C_3^2, σ_h
	$6/m$	C_{6h}	$C_6 \otimes C_i$
	622	D_6	$C_6 \otimes C_2(C_{2a})$
	$6mm$	C_{6v}	$C_6 \otimes C_s(\sigma_a)$
	$\bar{6}m2$	D_{3h}	$C_{3h} \otimes C_2(C_{2a})$
	$6/mmm$	D_{6h}	$D_6 \otimes C_i$
Cubic	23	T	E, C_{2m}, C_{3j}, C_{3j}^2
	$m\bar{3}$	T_h	$T \otimes C_i$
	432	O	E, C_{2m}, C_{3j}, C_{3j}^2, C_{2p}, C_{4m}, C_{4m}^3
	$\bar{4}3m$	T_d	E, C_{2m}, C_{3j}, C_{3j}^2, σ_p, S_{4m}, S_{4m}^3
	$m\bar{3}m$	O_h	$O \otimes C_i$

restrictions have to be introduced on the group of rotations and reflections that transform a Bravais lattice into itself.

Whenever the site at $\boldsymbol{R}_n$ is part of the lattice, so is the site at $-\boldsymbol{R}_n$. This implies that only point groups that contain inversion need to be taken into account. In the two-dimensional case 180° rotation (which is equivalent to inversion) has to be present.

A further restriction arises from the requirement that if there is an n-fold symmetry axis (with $n > 2$) among the symmetry elements then there must also be n mirror planes each of which contains the axis. In other words, among all point groups only those can belong to Bravais lattices for which the following is true: if the point group contains n (C_n) as a subgroup, then in the $n > 2$ case it should also contain the elements of the group nmm (for n even) or nm (C_{nv}) (for n odd). The point groups satisfying these conditions are called *Bravais groups*.

Thus in the two-dimensional case a rotation axis with $n > 2$ must be complemented by n mirror lines perpendicular to it. This is readily seen in Fig. 5.15. When the primitive vector $\boldsymbol{a}$ is rotated through $2\pi/n$ and $-2\pi/n$, each of the rotated vectors can be chosen as the second primitive vector provided $n > 2$. As the vector endpoints are lattice sites and the two points reached by the two rotations are each other's mirror images in the x-axis, this axis is also a mirror line for the entire lattice.

Four of the previous ten point groups meet these requirements:

$$2\,,\ \ 2mm\,,\ \ 4mm\,,\ \ 6mm\,.$$

The symmetries of the Bravais lattices of two-dimensional crystals are therefore given by one of these point groups.

The proof of the statement proceeds along similar lines in three dimensions. In this case the point group of the crystal's Bravais lattice must be one of seven point groups:

$$\bar{1}\,,\ \ 2/m\,,\ \ mmm\,,\ \ 4/mmm\,,\ \ \bar{3}m\,,\ \ 6/mmm\,,\ \ m\bar{3}m\,,$$

or, in Schoenflies notation,

$$S_2\,,\ \ C_{2h}\,,\ \ D_{2h}\,,\ \ D_{4h}\,,\ \ D_{3d}\,,\ \ D_{6h}\,,\ \ O_h\,.$$

Thus there are four two-dimensional and seven three-dimensional Bravais groups.

Two crystals are said to belong to the same *crystal system* or *syngony* if the point groups of their Bravais lattices are identical. This means that three-dimensional (two-dimensional) crystals can be grouped into seven (four) crystal systems.

Whenever possible, the primitive translation vectors of the Bravais lattice are chosen along a mirror line (in two dimensions) or a rotation axis. This choice provides the coordinate axes of the crystal system. The nomenclature of crystal systems is based on the relative orientation and possible equivalence of these axes under suitable symmetry transformations.

The four crystal systems of the two-dimensional case can be described as:

1. *Oblique* (m). Its symmetry group is the point group 2. There are no preferred directions within the plane. The two axes are equivalent but not perpendicular.

2. *Rectangular* (o). Its symmetry group is the point group $2mm$. There are two perpendicular mirror lines within the plane. The two crystallographic axes are not equivalent but they are perpendicular.
3. *Square* (t). Its symmetry group is the point group $4mm$. The perpendicular mirror lines possess a fourfold rotational symmetry. The two axes are equivalent and perpendicular.
4. *Hexagonal* (h). Its symmetry group is the point group $6mm$. Within the plane, equivalent mirror lines are at 120°. There are three equivalent crystallographic axes, making angles of $2\pi/3$.

The relations among the axes are shown in Fig. 5.17 for the three-dimensional crystal systems. They can be described as:

1. *Triclinic* (a). The point group of the underlying Bravais lattice is $\bar{1}$ (S_2). The three crystallographic axes are neither equivalent nor perpendicular.
2. *Monoclinic* (m). The point group of the underlying Bravais lattice is $2/m$ (C_{2h}). The three crystallographic axes are not equivalent but one is perpendicular to the plane spanned by the others.
3. *Orthorhombic* (o). The point group of the underlying Bravais lattice is mmm (D_{2h}). The three inequivalent crystallographic axes are mutually perpendicular.
4. *Tetragonal* (t). The point group of the underlying Bravais lattice is $4/mmm$ (D_{4h}). It has a fourfold principal axis and two equivalent secondary axes of equal length in the plane perpendicular to the principal axis.
5. *Rhombohedral* (trigonal, h). The point group of the underlying Bravais lattice is $\bar{3}m$ (D_{3d}). It has a threefold principal axis and three secondary axes of equal length in the plane perpendicular to the principal axis. However, neither the principal axis nor the secondary axes are along the primitive translation vectors. The three primitive vectors of the Bravais lattice are symmetric about the principal axis.
6. *Hexagonal* (h). The point group of the underlying Bravais lattice is $6/mmm$ (D_{6h}). It has a sixfold principal axis and three secondary axes of equal length that make angles of 120° with each other in the plane perpendicular to the principal axis. The relative orientations of the coordinate axes are therefore the same as in the trigonal system, justifying the common notation (h).[21]
7. *Cubic* (isometric, tesseral, c). The point group of the underlying Bravais lattice is $m\bar{3}m$ (O_h). Its has three equivalent axes that are mutually perpendicular.

[21] Crystallographers consider the rhombohedral and hexagonal crystal systems to belong to the same crystal family.

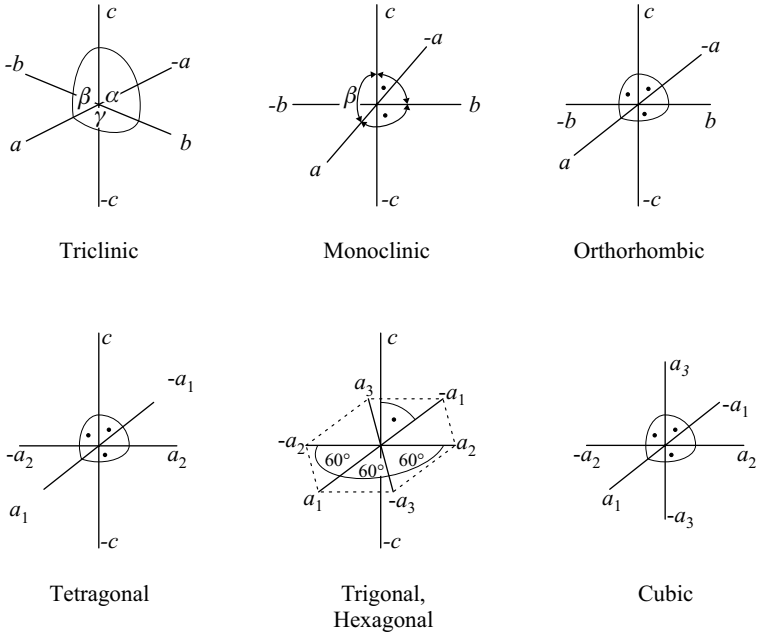


Fig. 5.17. Coordinate axes in the three-dimensional crystal systems

5.4.4 Two-Dimensional Bravais-Lattice Types

We shall now examine, on planar lattices first, the restrictions imposed on the primitive vectors by the symmetry operations of the point group that transform the lattice into itself.

A two-dimensional Bravais lattice is unambiguously determined by its two primitive vectors. In the most general case the primitive vectors are of unequal length and their angle φ is arbitrary. Without further restrictions the primitive cell is a rhomboid. Even then, the lattice has a twofold rotation axis that is perpendicular to the plane – thus its Bravais group is the point group 2 (C_2). Such a lattice belongs to the oblique crystal system. Since this classification is independent of the lengths and angles of the primitive translation vectors, lattices that belong to the same oblique system but have different lattice parameters are said to be of the same type. Formulated more generally: *Two Bravais lattices are said to be of the same type if they can be deformed into one another by continuously changing the lattice parameters (i.e., the lengths and angle of the primitive vectors) in such a way that the rotational and reflection symmetries of the lattice are conserved at all times.*

The only oblique lattice type is denoted by *mp*. Here *m* refers to the monoclinic lattice system, and *p* to *primitive*, as compared to *centered* lattices, which will be presented later.

The point group of lattices that belong to the rectangular crystal system is $2mm$. In addition to the twofold axis this point group contains two mutually perpendicular mirror lines. If the origin of coordinates is chosen at a lattice point and the axes along mirror lines, the primitive vectors can be written as $\boldsymbol{a}_i = a_i(\cos\varphi_i\,,\,\sin\varphi_i\,,\,0)$. The shortest of them is called $\boldsymbol{a}_1$. This is then reflected in the x-axis; its mirror image, $\boldsymbol{a}_1' = a_1(\cos\varphi_1\,,\,-\sin\varphi_1\,,\,0)$, as well as vectors

$$\boldsymbol{a}_1 - \boldsymbol{a}_1' = (0\,,\,2a_1\sin\varphi_1\,,\,0) \qquad \text{and} \qquad \boldsymbol{a}_1 + \boldsymbol{a}_1' = (2a_1\cos\varphi_1\,,\,0\,,\,0)$$

have to point to lattice sites. This requirement is met in three cases:

1. $\varphi_1 = 0$, and thus $\boldsymbol{a}_1' = \boldsymbol{a}_1$;
2. $\varphi_1 = 90°$, and thus $\boldsymbol{a}_1' = -\boldsymbol{a}_1$;
3. φ_1 is arbitrary, $\boldsymbol{a}_1'$ is linearly independent of $\boldsymbol{a}_1$, and thus $\boldsymbol{a}_1'$ is also a primitive vector. Reflection symmetry then implies that the two lattice vectors are of equal length.

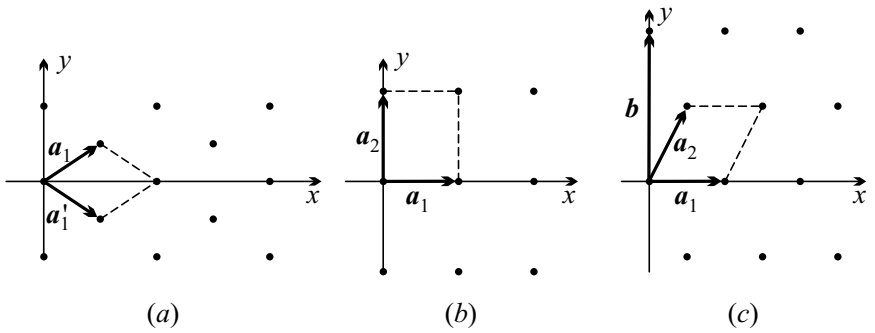


Fig. 5.18. Possible orientations of primitive vectors in planar lattices featuring the symmetries of the point group $2mm$

Consider the third case first. The orientation of primitive vectors is shown in Fig. 5.18(a). Since $\boldsymbol{a}_1$ and $\boldsymbol{a}_1'$ are of equal length, $\boldsymbol{a} = \boldsymbol{a}_1 + \boldsymbol{a}_1'$ and $\boldsymbol{b} = \boldsymbol{a}_1 - \boldsymbol{a}_1'$ are perpendicular and along the two mirror lines. $\boldsymbol{a}_1$ (as well as its mirror image) is the shortest among lattice vectors provided that φ is between 30° and 60°.

There is an equivalence between the first two cases – in which the shortest lattice vector is along or perpendicular to the selected mirror line –, since there are two mutually perpendicular mirror lines among the symmetry elements. Whichever possibility is chosen, the shortest lattice vector is along a symmetry line. Now consider the mirror image of the other primitive vector in the x-axis. As above, there are two possibilities, shown in parts (b) and (c) of Fig. 5.18. The second primitive vector is either along the other mirror line – in which case the two primitive vectors are perpendicular – or it is not at right angles

to the mirror lines – in which case the sum of $\boldsymbol{a}_2$ and its mirror image has to be $\boldsymbol{a}_1$. This is equivalent to saying that $\boldsymbol{b} = 2\boldsymbol{a}_2 - \boldsymbol{a}_1$ is perpendicular to $\boldsymbol{a}_1$.

The three lattices shown in Fig. 5.18 correspond in reality to two lattice types, as lattice (a) can be transformed into lattice (c) by an expansion in the y-direction and a compression in the x-direction. Throughout this procedure, the mirror lines remain unaltered. However, lattice (b) cannot be transformed into the other lattices without breaking the symmetry. This lattice type (b) is called *primitive rectangular* (op), and the others, (a) and (c) – *centered rectangular* (oc).

To understand the origin of the name note that the primitive vectors are along the mirror lines in lattice type (b), whereby the rectangular unit cell (primitive cell) also possesses the symmetries $2mm$ of the Bravais lattice. On the other hand, the primitive cell is a parallelogram for the lattice type shown in parts (a) and (c), which does not exhibit the symmetries of the point group $2mm$. The rectangle spanned by the vectors $\boldsymbol{a} = \boldsymbol{a}_1 + \boldsymbol{a}_2$ and $\boldsymbol{b} = \boldsymbol{a}_1 - \boldsymbol{a}_2$ (or $\boldsymbol{a} = \boldsymbol{a}_1$ and $\boldsymbol{b} = 2\boldsymbol{a}_2 - \boldsymbol{a}_1$) possesses the symmetry of the point group of the Bravais lattice – but it is not a primitive cell as the center of the rectangle is also a lattice point. Nevertheless it is more practical to span the lattice using vectors $\boldsymbol{a}$ and $\boldsymbol{b}$, as it has already been mentioned in connection with Fig. 5.9. When this choice is made, a lattice point has to be included at the center of each rectangle – hence the name *centered rectangular lattice*. The rectangular nonprimitive cell spanned by the vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ is a conventional unit cell also called the *Bravais cell*. In general, if the primitive cell does not exhibit the point-group symmetry of the lattice but it is possible to choose a larger Bravais cell that does, then the lattice type is called centered because the Bravais cell contains one or more further lattice points at high-symmetry positions. Otherwise the lattice type is called simple or primitive.

Finally, we have to consider lattices that show the symmetries of the point groups $4mm$ and $6mm$. Because of the four- and sixfold rotational symmetry, the primitive lattice vectors are of equal length in both cases. When the rotation axis is fourfold, the two primitive vectors are also perpendicular, so the primitive cell is a square, hence the name square crystal system. The length of the side of the square can be arbitrary, thus the Bravais lattice of any crystal in the crystal system is of the same primitive square type, tp. When the rotation axis is sixfold, the equally long primitive vectors can make angles of $60°$ or $120°$. Here, again, the Bravais lattice of each crystal that belongs to the hexagonal system is of the same type, hp.

To summarize, two-dimensional crystal structures are divided into four systems according to the point group of their underlying Bravais lattices: oblique (m), rectangular (o), square (t), and hexagonal (h) – however, five Bravais lattice types are distinguished, since primitive and centered lattice types are equally possible within the rectangular system. These lattice types are shown in Fig. 5.19. The relations between the parameters of the Bravais lattices are listed in Table 5.6.

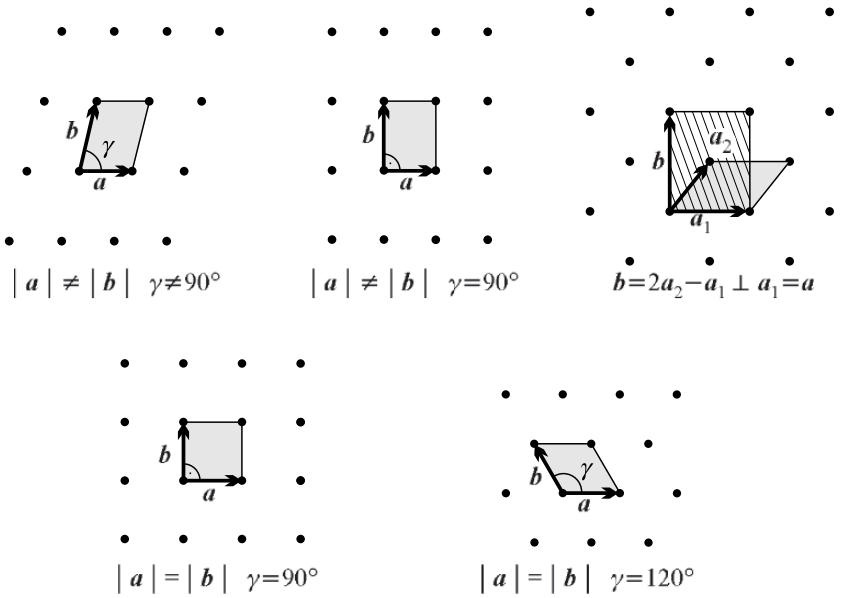


Fig. 5.19. Types of Bravais lattices and their primitive cells in two-dimensional crystal structures. The conventional unit cell (Bravais cell) is also indicated for the centered rectangular lattice

Table 5.6. The point groups for the four two-dimensional crystal systems (five Bravais lattice types), and relations for the parameters of the Bravais cell

System	Type	Symbol	Point group	Lattice parameters
Oblique	primitive	mp	2	$a \neq b\,,\ \gamma \neq 90^\circ$
Rectangular	primitive centered	op oc	$2mm$	$a \neq b\,,\ \gamma = 90^\circ$
Square	primitive	tp	$4mm$	$a = b\,,\ \gamma = 90^\circ$
Hexagonal	primitive	hp	$6mm$	$a = b\,,\ \gamma = 120^\circ$

It has already been mentioned that one possible choice for the primitive cell is the Wigner–Seitz cell. This is shown for each of the five planar lattice types in Fig. 5.20.

It is readily seen for the hexagonal Bravais lattice that the Wigner–Seitz cell has a much higher symmetry than the primitive cell spanned by the primitive translation vectors. In three dimensional Bravais lattices we shall see further examples in which the Wigner–Seitz cell has a much more complicated form than parallelepiped-shaped primitive cells but it also has a higher symmetry. This property fully justifies its usage.

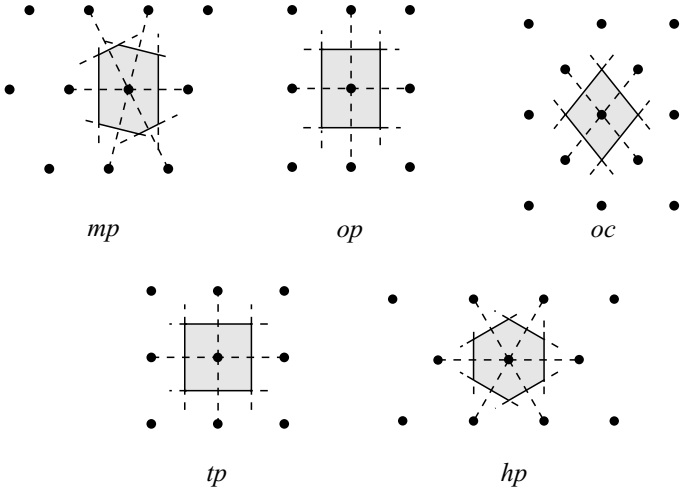


Fig. 5.20. Wigner–Seitz cells for the five planar lattice types

5.4.5 Three-Dimensional Bravais-Lattice Types

It is straightforward to generalize the results of the two-dimensional case to three dimensions. However, this would be rather lengthy, therefore on many occasions we shall content ourselves with presenting the most important results without derivation.

Corresponding to the seven crystal systems (syngonies) there are seven primitive (P) Bravais-lattice types. With the exception of the rhombohedral system, the primitive translation vectors of these lattices are along the directions that correspond to symmetries of the Bravais group. When and only when two axes are equivalent, the length of the associated primitive vectors are equal. The simplest case is that of the Bravais group $\bar{1}$ (S_2), whose only element other than the identity is inversion. This symmetry places no restriction on the lattice parameters: the lengths a_1, a_2, a_3 of the primitive translation vectors $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$, as well as their angles α_1, α_2, α_3 are arbitrary.

In lattices of higher symmetry, symmetry elements provide relations among the axes – and thereby also among the lengths and/or orientation of the primitive vectors. As we shall see, just like in the two-dimensional case, there are some three-dimensional Bravais lattices in which the primitive cell does not show the symmetry of the underlying lattice but a larger Bravais cell can be chosen that exhibits the point-group symmetry of the lattice. The Bravais cell is spanned by the vectors $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$, which are appropriate linear combinations of the primitive translation vectors $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$. Note that in what follows we shall always use $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, and $\boldsymbol{a}_3$ for primitive vectors, $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ for the edge vectors of the Bravais cell, and α, β, γ for the angles of the latter. For primitive lattice types the two cells are identical, and so the two notations refer to the same trio of basis vectors.

Symmetry-implied restrictions on the side lengths and angles of the Bravais cell are listed in Table 5.7. The name of the crystal system refers to the shape of the Bravais cell.

Table 5.7. Restrictions on the lattice parameters of the Bravais cell and characteristic symmetry operations for lattices of the seven crystal systems

System	Symbol	Lattice parameters	Characteristic symmetry
Triclinic	a	$a \neq b \neq c$ $\alpha \neq \beta \neq \gamma$	inversion only
Monoclinic	m	$a \neq b \neq c$ $\alpha = \gamma = 90^\circ \neq \beta$	rotation through 180° around a single axis
Orthorhombic	o	$a \neq b \neq c$ $\alpha = \beta = \gamma = 90^\circ$	rotation through 180° around three mutually perpendicular axes
Tetragonal	t	$a = b \neq c$ $\alpha = \beta = \gamma = 90^\circ$	rotation through 90° around a single axis
Rhombohedral (Trigonal)	h	$a = b = c$ $120^\circ > \alpha = \beta = \gamma \neq 90^\circ$	rotation through 120° around a single axis
Hexagonal	h	$a = b \neq c$ $\alpha = \beta = 90^\circ$, $\gamma = 120^\circ$	rotation through 60° around a single axis
Cubic	c	$a = b = c$ $\alpha = \beta = \gamma = 90^\circ$	rotation through 120° around the four space diagonals

The crystal system specified by a particular Bravais group and the relationship between the lattice parameters of the Bravais cell may lead to various Bravais-lattice types in the three-dimensional case, too. We shall now illustrate this by two examples. Figure 5.21(a) shows the primitive cell of an orthorhombic Bravais lattice spanned by three mutually perpendicular primitive vectors of different length. The cuboid (rectangular parallelepiped) shaped primitive cell possesses the characteristic symmetries of the orthorhombic crystal system: rotations through 180° around three mutually perpendicular axes.

Among the primitive translation vectors of the Bravais lattice shown in part (b) of Fig. 5.21, $\boldsymbol{a}_3$ is perpendicular to the plane spanned by $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$. The latter are not perpendicular to one another but satisfy the condition $(2\boldsymbol{a}_1 - \boldsymbol{a}_2) \perp \boldsymbol{a}_2$. We shall now prove that even in this case the lattice possesses three mutually perpendicular twofold axes, and therefore belongs to the orthorhombic crystal system. To this end consider the cell spanned by the translation vectors $\boldsymbol{a} = 2\boldsymbol{a}_1 - \boldsymbol{a}_2$, $\boldsymbol{b} = \boldsymbol{a}_2$, and $\boldsymbol{c} = \boldsymbol{a}_3$. The condition imposed on the primitive vectors implies that this cell is a cuboid. This unit cell is

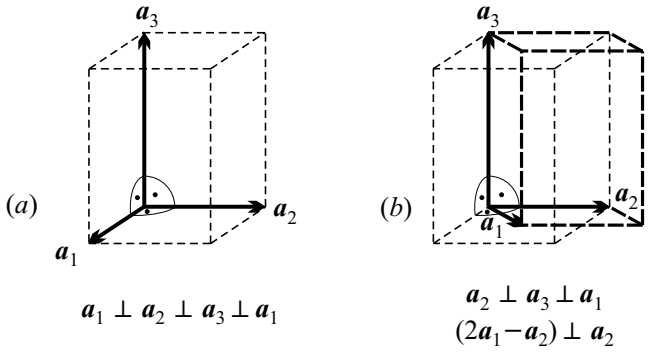


Fig. 5.21. Bravais-lattice types of the orthorhombic crystal system: (*a*) primitive; (*b*) base-centered

not primitive, as it contains two lattice points: one at a vertex and another at the center of a face that we shall refer to as of the base. (The equivalent opposite face is also centered but it belongs to the neighboring primitive cell.) Nevertheless, it is often more convenient to work with it. When the directions of the vectors $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ are chosen as axes, the coordinate axes of the orthorhombic system are recovered, and so the relations among the parameters given in Table 5.7 are satisfied by the Bravais cell spanned by these vectors. Because of the lattice point at the center of a face (base) this lattice is called *single-face-centered orthorhombic* or more commonly *base-centered orthorhombic*. This lattice type is denoted by C – or A or B when another face (that is not perpendicular to $\boldsymbol{c}$) is centered.[22]

As a second example, consider lattices of the cubic crystal system that are invariant under the symmetries of the point group O_h. To obtain such lattices, the edge vectors of the primitive cell need not be mutually perpendicular and of equal length. In general, when the three primitive translation vectors are of equal length and their angles are equal but not right angles, the symmetry is trigonal. The threefold axis is the resultant of the three primitive vectors. However, the three vectors may happen to make angles $\varphi = \arccos(-1/3)$ or $\varphi = \arccos(1/2)$ with one another. As shown in Fig. 5.22, by choosing

$$\boldsymbol{a} = \boldsymbol{a}_2 + \boldsymbol{a}_3\,, \qquad \boldsymbol{b} = \boldsymbol{a}_1 + \boldsymbol{a}_3\,, \qquad \boldsymbol{c} = \boldsymbol{a}_1 + \boldsymbol{a}_2 \tag{5.4.5}$$

in the first case and

$$\boldsymbol{a} = -\boldsymbol{a}_1 + \boldsymbol{a}_2 + \boldsymbol{a}_3\,, \quad \boldsymbol{b} = \boldsymbol{a}_1 - \boldsymbol{a}_2 + \boldsymbol{a}_3\,, \quad \boldsymbol{c} = \boldsymbol{a}_1 + \boldsymbol{a}_2 - \boldsymbol{a}_3 \tag{5.4.6}$$

in the second, one is left with three mutually perpendicular vectors $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ of the same length forming a cube. The primitive translation vectors are now along face or body diagonals of this cube.

[22] Although most books follow this convention, the use of letter S derived from "side-face-centered" has been recommended recently by a committee of the International Union of Crystallography.

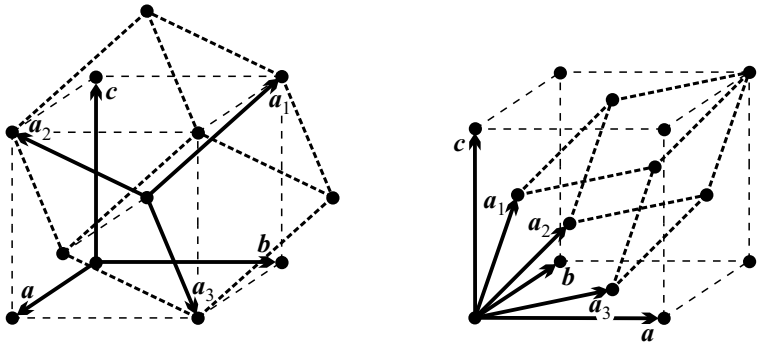


Fig. 5.22. Primitive translation vectors and Bravais cells for body-centered and face-centered cubic lattices

In addition to the vertices, the cubic Bravais cells or conventional unit cells now contain lattice sites either at the body centers or the face centers. These lattices are considered to belong to different lattice types because they cannot be continuously deformed into one another or into a simple cubic lattice without breaking some symmetries. For example, if the length of the primitive vectors were left unaltered but the relative angles were changed, only symmetries typical of the trigonal system could be conserved upon the smallest deformation. The lattice type that contains a site at the center of the Bravais cell is called *body-centered*; its symbol is I. The lattice type that contains a site at each face center is called *face-centered*; its symbol is F.

From the above examples one may expect that new lattice types of the same crystal system can be obtained by adding further sites at special positions of the primitive cell – at the body center, at the centers of a pair of opposite faces, or at the center of each face provided the point-group symmetry is unchanged. More general locations are certainly excluded since all the sites of the original lattice have to be reached by integer linear combinations of the new, shorter primitive translation vectors.

The seven crystal systems and the three types of centering could thus be expected to give rise to 21 types of centered lattices. In reality, only seven centered lattice types appear as centering does not always lead to new lattice types: in particular, in the triclinic system centering does not lead to a single new type. In other cases symmetries of the simple lattice are broken in the centered lattice. The new lattice types and the primitive ones are listed Table 5.8 for each crystal system.

The variety of notational conventions used for point groups exists for Bravais lattice types, too. In the Schoenflies notation the subscript of Γ specifies the crystal system (t = triclinic, m = monoclinic, o = orthorhombic, q = tetragonal, rh = rhombohedral, h = hexagonal, c = cubic), and the superscript refers to the centering the Bravais cell (c = base-centered, v = body-centered, f = face-centered). Another convention uses a code of the form xY,

Table 5.8. The seven crystal systems and the fourteen types of Bravais lattices

Crystal system		Point group		Type				
Name	Symbol			P	C	I	F	R
Triclinic	a	$\bar{1}$	C_i	$\Gamma_{\rm t}$				
Monoclinic	m	$2/m$	C_{2h}	$\Gamma_{\rm m}$	$\Gamma_{\rm m}^{\rm c}$			
Orthorhombic	o	mmm	D_{2h}	$\Gamma_{\rm o}$	$\Gamma_{\rm o}^{\rm c}$	$\Gamma_{\rm o}^{\rm v}$	$\Gamma_{\rm o}^{\rm f}$	
Tetragonal	t	$4/mmm$	D_{4h}	$\Gamma_{\rm q}$		$\Gamma_{\rm q}^{\rm v}$		
Rhombohedral	h	$\bar{3}m$	D_{3d}					$\Gamma_{\rm rh}$
Hexagonal	h	$6/mmm$	D_{6h}	$\Gamma_{\rm h}$				
Cubic	c	$m\bar{3}m$	O_h	$\Gamma_{\rm c}$		$\Gamma_{\rm c}^{\rm v}$	$\Gamma_{\rm c}^{\rm f}$	

where x is the international symbol for the crystal system (2nd column), and Y – one of the letters P, C (S), I, or F – specifies the centering type of the lattice. The rhombohedral (trigonal) lattice is an exception: although it is a primitive lattice, its centering type is denoted by R. Thus hP stands for hexagonal, and hR for rhombohedral crystal system. The reader will understand in hindsight why the rhombohedral system appears as a nonprimitive type of the hexagonal system. A third notation uses a code of the form Yx, where Y, once again, refers to the centering type ($Y = P$, C, I, F, R), and x is the short international symbol for the point group of the Bravais lattice. E.g., for body-centered cubic lattices the notations $\Gamma_{\rm c}^{\rm v}$, cI, and $Im\bar{3}m$ are used equally. All three notations are given in parentheses at the listing of Bravais-lattice types.

Below we list the fourteen Bravais-lattice types, together with the relations among the vectors spanning the primitive cell and the Bravais cell. (Bravais cells are shown in Fig. 5.23.) The two triplets of vectors are the same for primitive lattices but not for nonprimitive ones. As mentioned at the introduction of primitive translation vectors on page 111, there is some arbitrariness in their choice. That is why the choice of vectors presented below is not the only one used in the literature.

1. *Triclinic* (aP, $P\bar{1}$, $\Gamma_{\rm t}$). The three basis vectors ($\boldsymbol{a} = \boldsymbol{a}_1$, $\boldsymbol{b} = \boldsymbol{a}_2$, $\boldsymbol{c} = \boldsymbol{a}_3$) can point to any directions, and their lengths are arbitrary, too: $a \neq b \neq c$, $\alpha \neq \beta \neq \gamma$.
2. *Simple monoclinic* (primitive monoclinic) (mP, $P2/m$, $\Gamma_{\rm m}$). The lengths of the three basis vectors are arbitrary, $a \neq b \neq c$, however one of them (customarily chosen as $\boldsymbol{b}$) is perpendicular to the plane spanned by the others, $\alpha = \gamma = 90° \neq \beta$. In rectangular coordinates the primitive vectors are given by $\boldsymbol{a} = \boldsymbol{a}_1 = (a\sin\beta, 0, a\cos\beta)$, $\boldsymbol{b} = \boldsymbol{a}_2 = (0, b, 0)$, $\boldsymbol{c} = \boldsymbol{a}_3 = (0, 0, c)$.
3. *Centered monoclinic* (base-centered monoclinic) (mC or mS, $C2/m$, $\Gamma_{\rm m}^{\rm b}$). A lattice site is added at the center of either rectangular face (base) of the simple monoclinic Bravais lattice. When the vectors specified in the

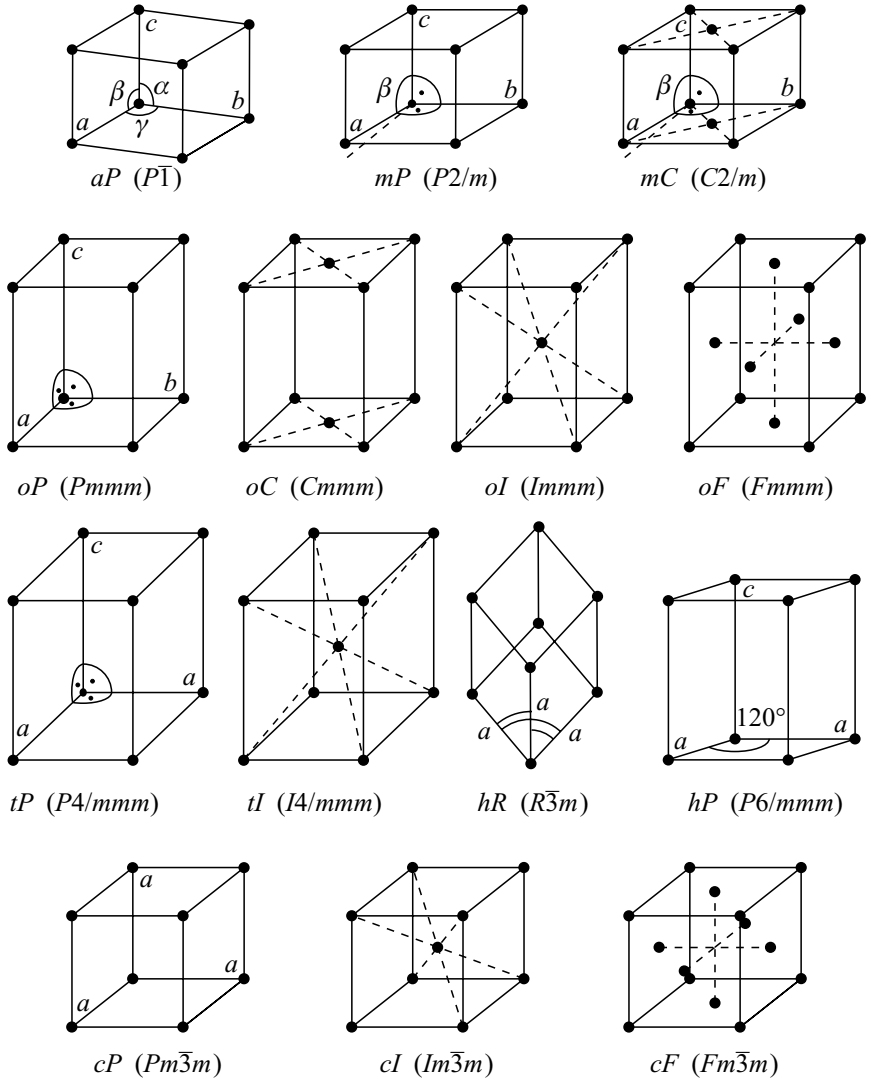


Fig. 5.23. The fourteen types of three-dimensional Bravais lattices

simple monoclinic case are chosen as the edge vectors of the Bravais cell of sides a, b, c, i.e., $\boldsymbol{a} = (a\sin\beta, 0, a\cos\beta)$, $\boldsymbol{b} = (0, b, 0)$, $\boldsymbol{c} = (0, 0, c)$, then the primitive translation vectors may be chosen as $\boldsymbol{a}_1 = \boldsymbol{a}$, $\boldsymbol{a}_2 = \frac{1}{2}(\boldsymbol{a} + \boldsymbol{b})$, $\boldsymbol{a}_3 = \boldsymbol{c}$. The three primitive vectors are of unequal length, and neither two are perpendicular. Nevertheless the lattice is not triclinic, as the orientation of the vectors is such that $2\boldsymbol{a}_2 - \boldsymbol{a}_1$ is perpendicular to the plane spanned by $\boldsymbol{a}_1$ and $\boldsymbol{a}_3$. This property gives rise to a twofold symmetry axis along the direction $\boldsymbol{b} = 2\boldsymbol{a}_2 - \boldsymbol{a}_1$.

Note that the base-centered monoclinic Bravais lattice can also be considered as a body-centered monoclinic Bravais lattice, as it is sometimes done in the literature. When $2\boldsymbol{a}_2 - \boldsymbol{a}_1$ is perpendicular to the plane spanned by $\boldsymbol{a}_1$ and $\boldsymbol{a}_3$ then the choice $\boldsymbol{a} = \boldsymbol{a}_1 - \boldsymbol{a}_3$, $\boldsymbol{b} = 2\boldsymbol{a}_2 - \boldsymbol{a}_1$, $\boldsymbol{c} = \boldsymbol{a}_3$ leads to a monoclinic Bravais cell that has a lattice site at its body center.

4. *Simple orthorhombic* (primitive orthorhombic) (oP, $Pmmm$, Γ_{o}). The lengths of the three basis vectors are arbitrary, $a \neq b \neq c$, however they are mutually perpendicular: $\boldsymbol{a}_1 = \boldsymbol{a} \perp \boldsymbol{a}_2 = \boldsymbol{b} \perp \boldsymbol{a}_3 = \boldsymbol{c} \perp \boldsymbol{a}_1$. In rectangular coordinates, $\boldsymbol{a}_1 = (a, 0, 0)$, $\boldsymbol{a}_2 = (0, b, 0)$, $\boldsymbol{a}_3 = (0, 0, c)$. The Bravais cell is the same as the primitive cell. $a \neq b \neq c$, $\alpha = \beta = \gamma = 90°$.
5. *Single-face-centered orthorhombic* (base-centered orthorhombic) (oC or oS, $Cmmm$, $\Gamma_{\mathrm{o}}^{\mathrm{b}}$). The Bravais cell has the characteristic properties of a simple orthorhombic Bravais lattice. Its sides and angles satisfy the relations $a \neq b \neq c$, $\alpha = \beta = \gamma = 90°$. However, an additional lattice site is found at the centers of two equivalent opposite faces of the Bravais cell. This implies that the primitive translation vectors are $\boldsymbol{a}_1 = \frac{1}{2}(a, -b, 0)$, $\boldsymbol{a}_2 = \frac{1}{2}(a, b, 0)$, and $\boldsymbol{a}_3 = (0, 0, c)$. Vectors $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ are not perpendicular to one another but $\boldsymbol{a}_3$ is perpendicular to them – therefore the three vectors seemingly form a monoclinic system. If, however, $\boldsymbol{a}_1 + \boldsymbol{a}_2$ and $\boldsymbol{a}_2 - \boldsymbol{a}_1$ are also perpendicular then additional symmetries appear, resulting in a symmetry group mmm (D_{2h}) instead of $2/m$ (C_{2h}).
6. *Body-centered orthorhombic* (oI, $Immm$, $\Gamma_{\mathrm{o}}^{\mathrm{v}}$). The sides and angles of the Bravais cell once again satisfy the relations $a \neq b \neq c$, $\alpha = \beta = \gamma = 90°$, however the additional lattice site is now found at the center of the cell. A customary choice for the primitive translation vectors is $\boldsymbol{a}_1 = \frac{1}{2}(-a, b, c)$, $\boldsymbol{a}_2 = \frac{1}{2}(a, -b, c)$, $\boldsymbol{a}_3 = \frac{1}{2}(a, b, -c)$. These are related to the edge vectors of the Bravais cell through $\boldsymbol{a} = \boldsymbol{a}_2 + \boldsymbol{a}_3$, $\boldsymbol{b} = \boldsymbol{a}_1 + \boldsymbol{a}_3$, $\boldsymbol{c} = \boldsymbol{a}_1 + \boldsymbol{a}_2$. Note that it is equally possible to choose two primitive vectors along the sides of the Bravais cell, $\boldsymbol{a}_1 = \boldsymbol{a} = (a, 0, 0)$, $\boldsymbol{a}_2 = \boldsymbol{b} = (0, b, 0)$; in this case the third has to satisfy $\boldsymbol{c} = 2\boldsymbol{a}_3 - (\boldsymbol{a}_1 + \boldsymbol{a}_2)$ – that is, $\boldsymbol{a}_3 = \frac{1}{2}(a, b, c)$.
7. *All-faces-centered orthorhombic* (face-centered orthorhombic) (oF, $Fmmm$, $\Gamma_{\mathrm{o}}^{\mathrm{f}}$). The Bravais cell is the same as in other orthorhombic Bravais lattices, $a \neq b \neq c$, $\alpha = \beta = \gamma = 90°$, however additional lattice sites are now found at each face center. The customary choice for the primitive translation vectors is $\boldsymbol{a}_1 = \frac{1}{2}(0, b, c)$, $\boldsymbol{a}_2 = \frac{1}{2}(a, 0, c)$, $\boldsymbol{a}_3 = \frac{1}{2}(a, b, 0)$. These are related to the edge vectors of the Bravais cell through $\boldsymbol{a} = -\boldsymbol{a}_1 + \boldsymbol{a}_2 + \boldsymbol{a}_3$, $\boldsymbol{b} = \boldsymbol{a}_1 - \boldsymbol{a}_2 + \boldsymbol{a}_3$, $\boldsymbol{c} = \boldsymbol{a}_1 + \boldsymbol{a}_2 - \boldsymbol{a}_3$. If a primitive vector is chosen to lie along the edge of the Bravais cell, $\boldsymbol{a}_1 = \boldsymbol{a} = (a, 0, 0)$, then the two others have to be chosen in such a manner that $\boldsymbol{b} = 2\boldsymbol{a}_2 - \boldsymbol{a}_1$ and $\boldsymbol{c} = 2\boldsymbol{a}_3 - \boldsymbol{a}_1$ are perpendicular to $\boldsymbol{a}_1$. This implies $\boldsymbol{a}_2 = \frac{1}{2}(a, b, 0)$, $\boldsymbol{a}_3 = \frac{1}{2}(a, 0, c)$.
8. *Simple tetragonal* (primitive tetragonal) (tP, $P4/mmm$, Γ_{q}). The three primitive vectors are mutually perpendicular, but only two of them are of equal length: $\boldsymbol{a}_1 \perp \boldsymbol{a}_2 \perp \boldsymbol{a}_3 \perp \boldsymbol{a}_1$, $|\boldsymbol{a}_1| = |\boldsymbol{a}_2| \neq |\boldsymbol{a}_3|$. The Bravais cell is the same as the primitive cell. $\boldsymbol{a} = \boldsymbol{a}_1 = (a, 0, 0)$, $\boldsymbol{b} = \boldsymbol{a}_2 = (0, a, 0)$, $\boldsymbol{c} = \boldsymbol{a}_3 = (0, 0, c)$, $a = b \neq c$, $\alpha = \beta = \gamma = 90°$.

9. *Body-centered tetragonal* (tI, $I4/mmm$, $\Gamma_{\mathrm{q}}^{\mathrm{v}}$). To obtain a new type of Bravais lattice from the simple tetragonal lattice, one has to add a lattice site at the body center of each Bravais cell. The relations $a = b \neq c$, $\alpha = \beta = \gamma = 90°$ continue to hold for the Bravais cell, however the primitive vectors can now be chosen as $\boldsymbol{a}_1 = \frac{1}{2}(-a, a, c)$, $\boldsymbol{a}_2 = \frac{1}{2}(a, -a, c)$, $\boldsymbol{a}_3 = \frac{1}{2}(a, a, -c)$. Similarly to the case of the body-centered orthorhombic lattice, the primitive vectors are related to the edge vectors of the Bravais cell by $\boldsymbol{a} = \boldsymbol{a}_2 + \boldsymbol{a}_3$, $\boldsymbol{b} = \boldsymbol{a}_1 + \boldsymbol{a}_3$, $\boldsymbol{c} = \boldsymbol{a}_1 + \boldsymbol{a}_2$. Once again, it is possible to choose two primitive translation vectors along the sides of the Bravais cell, e.g., $\boldsymbol{a}_1 = \boldsymbol{a} = (a, 0, 0)$, $\boldsymbol{a}_2 = \boldsymbol{b} = (0, a, 0)$; in this case the third vector has to satisfy the condition $\boldsymbol{c} = 2\boldsymbol{a}_3 - (\boldsymbol{a}_1 + \boldsymbol{a}_2)$ – that is, $\boldsymbol{a}_3 = \frac{1}{2}(a, a, c)$.
10. *Rhombohedral* (hR, $R\bar{3}m$, Γ_{rh}). The three primitive vectors are of equal length and make equal (but not right) angles with each other. As the three vectors cannot be coplanar, their angles can be chosen to be smaller than 120° without any loss of generality. The choice $\boldsymbol{a} = \boldsymbol{a}_1$, $\boldsymbol{b} = \boldsymbol{a}_2$, $\boldsymbol{c} = \boldsymbol{a}_3$ implies $a = b = c$ and $120° > \alpha = \beta = \gamma \neq 90°$. In rectangular coordinates a customary choice for the primitive vectors is $\boldsymbol{a}_1 = (a, 0, c)$ $\boldsymbol{a}_2 = \frac{1}{2}(-a, \sqrt{3}a, 2c)$, $\boldsymbol{a}_3 = \frac{1}{2}(-a, -\sqrt{3}a, 2c)$.
11. *Simple hexagonal* (primitive hexagonal) (hP, $P6/mmm$, Γ_{h}). Two of the primitive vectors ($\boldsymbol{a}_1$ and $\boldsymbol{a}_2$) are of equal length and make an angle of 120°; the third one ($\boldsymbol{a}_3$) is perpendicular to them. The Bravais cell is the same as the primitive cell. A possible choice for the primitive vectors is $\boldsymbol{a} = \boldsymbol{a}_1 = \frac{1}{2}(\sqrt{3}a, -a, 0)$, $\boldsymbol{b} = \boldsymbol{a}_2 = (0, a, 0)$, $\boldsymbol{c} = \boldsymbol{a}_3 = (0, 0, c)$. $a = b \neq c$, $\alpha = \beta = 90°$, $\gamma = 120°$.
    The rhombohedral Bravais lattice may also be considered as a centered version of the hexagonal lattice. This is reflected by the notation hR. Two lattice sites need to be added to the Bravais cell of a hexagonal lattice, at $\frac{2}{3}\boldsymbol{a} + \frac{1}{3}\boldsymbol{b} + \frac{1}{3}\boldsymbol{c}$ and $\frac{1}{3}\boldsymbol{a} + \frac{2}{3}\boldsymbol{b} + \frac{2}{3}\boldsymbol{c}$. In rectangular coordinates these points are expressed as $\left(\frac{1}{3}\sqrt{3}a, 0, \frac{1}{3}c\right)$ and $\left(\frac{1}{6}\sqrt{3}a, \frac{1}{2}a, \frac{2}{3}c\right)$. The primitive vectors in this new lattice are thus

$$\boldsymbol{a}_1 = \left(\tfrac{1}{3}\sqrt{3}a, 0, \tfrac{1}{3}c\right), \quad \boldsymbol{a}_2 = \left(-\tfrac{1}{6}\sqrt{3}a, \tfrac{1}{2}a, \tfrac{1}{3}c\right), \quad \boldsymbol{a}_3 = \left(-\tfrac{1}{6}\sqrt{3}a, -\tfrac{1}{2}a, \tfrac{1}{3}c\right).$$

    These vectors are oriented the same way as the primitive vectors of a simple rhombohedral lattice. However, because of centering, this new lattice has a lower symmetry than the hexagonal one, which justifies considering it as a different lattice type.
12. *Simple cubic* (primitive cubic) (cP, $Pm\bar{3}m$, Γ_{c}). The three primitive vectors are of equal length and mutually perpendicular: $\boldsymbol{a}_1 \perp \boldsymbol{a}_2 \perp \boldsymbol{a}_3 \perp \boldsymbol{a}_1$ and $|\boldsymbol{a}_1| = |\boldsymbol{a}_2| = |\boldsymbol{a}_3|$. Here, too, the Bravais cell is the same as the primitive cell. $a = b = c$, $\alpha = \beta = \gamma = 90°$.
13. *Body-centered cubic* (bcc) (cI, $Im\bar{3}m$, $\Gamma_{\mathrm{c}}^{\mathrm{v}}$). As above, the Bravais cell is cubic, i.e., its parameters satisfy the conditions $a = b = c$, $\alpha = \beta = \gamma = 90°$, however an additional lattice point is found at the center of the cell. Thus the primitive vectors may be chosen as $\boldsymbol{a}_1 = \frac{1}{2}(-a, a, a)$, $\boldsymbol{a}_2 = \frac{1}{2}(a, -a, a)$, $\boldsymbol{a}_3 = \frac{1}{2}(a, a, -a)$. Just like in body-centered tetragonal

lattices, they are related to the edge vectors of the Bravais cell by $\boldsymbol{a} = \boldsymbol{a}_2 + \boldsymbol{a}_3$, $\boldsymbol{b} = \boldsymbol{a}_1 + \boldsymbol{a}_3$, $\boldsymbol{c} = \boldsymbol{a}_1 + \boldsymbol{a}_2$. As an alternative, two primitive vectors can again be chosen along the sides of the Bravais cell, $\boldsymbol{a}_1 = \boldsymbol{a} = (a, 0, 0)$, $\boldsymbol{a}_2 = \boldsymbol{b} = (0, a, 0)$, in which case the third must satisfy the requirement $\boldsymbol{c} = 2\boldsymbol{a}_3 - (\boldsymbol{a}_1 + \boldsymbol{a}_2)$ – that is, $\boldsymbol{a}_3 = \frac{1}{2}(a, a, a)$.

14. *Face-centered cubic* (fcc) (cF, $Fm\bar{3}m$, $\Gamma_{\mathrm{c}}^{\mathrm{f}}$). Once again, the lattice parameters of the Bravais cell satisfy the conditions $a = b = c$ and $\alpha = \beta = \gamma = 90°$, however additional lattice sites are found at each face center. As it has been mentioned, the customary choice for the primitive vectors is $\boldsymbol{a}_1 = \frac{1}{2}(0, a, a)$, $\boldsymbol{a}_2 = \frac{1}{2}(a, 0, a)$, $\boldsymbol{a}_3 = \frac{1}{2}(a, a, 0)$. These are related to the edge vectors of the Bravais cell by $\boldsymbol{a} = -\boldsymbol{a}_1 + \boldsymbol{a}_2 + \boldsymbol{a}_3$, $\boldsymbol{b} = \boldsymbol{a}_1 - \boldsymbol{a}_2 + \boldsymbol{a}_3$, $\boldsymbol{c} = \boldsymbol{a}_1 + \boldsymbol{a}_2 - \boldsymbol{a}_3$. Another common choice is taking two perpendicular vectors, $\boldsymbol{a}_1 = \frac{1}{2}(a, -a, 0)$ and $\boldsymbol{a}_2 = \frac{1}{2}(a, a, 0)$, and choosing the third one so that $2\boldsymbol{a}_3 - (\boldsymbol{a}_1 + \boldsymbol{a}_2)$ should be perpendicular to the plane of $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$. This requirement is satisfied by $\boldsymbol{a}_3 = \frac{1}{2}(a, 0, a)$. The three edge vectors of the Bravais cell are mutually perpendicular and of equal length: $\boldsymbol{a} = \boldsymbol{a}_1 + \boldsymbol{a}_2$, $\boldsymbol{b} = \boldsymbol{a}_2 - \boldsymbol{a}_1$, and $\boldsymbol{c} = 2\boldsymbol{a}_3 - (\boldsymbol{a}_1 + \boldsymbol{a}_2)$.

5.4.6 The Hierarchy of Crystal Systems

The previous enumeration of crystal systems proceeded from the simplest point groups to those with more symmetry elements. By adopting the opposite approach, a certain natural hierarchy can be observed among crystal systems. The most straightforward way to demonstrate this is to take a cubic (in two dimensions: square) or hexagonal crystal system, and to reduce the symmetry by appropriate deformations of the crystal.

Of all two-dimensional point lattices square and hexagonal ones possess the highest symmetries. Small deformations cannot take them into one another. A small deformation of the square lattice – stretching one of its sides – leads to a simple rectangular lattice, while stretching a square lattice along its diagonal leads to a centered rectangular lattice. The same lattice type is obtained when a hexagonal lattice is stretched or compressed along a mirror line or in a direction perpendicular to it. Square and hexagonal crystal systems are thus said to be higher in the hierarchy than the rectangular one. Further (shearing) deformation of both types of rectangular lattices leads to oblique lattices. The hierarchy of two-dimensional crystal systems is summarized in Fig. 5.24.

In the three-dimensional case we shall start off with the most regular primitive cell, the cubic one. It can be deformed by pulling or pushing on two opposite faces: angles are left intact while a side is stretched or compressed. The result is a rectangular prism with a square base – an object with tetragonal symmetry.

By stretching the object in another direction, a general rectangular parallelepiped with orthorhombic symmetry is obtained. A further shearing deformation along one of the planes changes the inclination of the edge initially

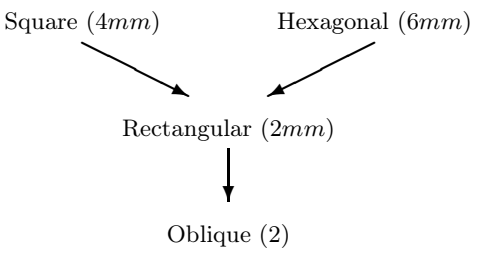


Fig. 5.24. Hierarchy of two-dimensional crystal systems

perpendicular to the shearing plane. The result is an object that possesses only monoclinic symmetry. Finally, another shearing in another plane leads to the most general triclinic parallelepiped. The objects obtained through the above sequence of deformations of a cube are shown in Fig. 5.25.

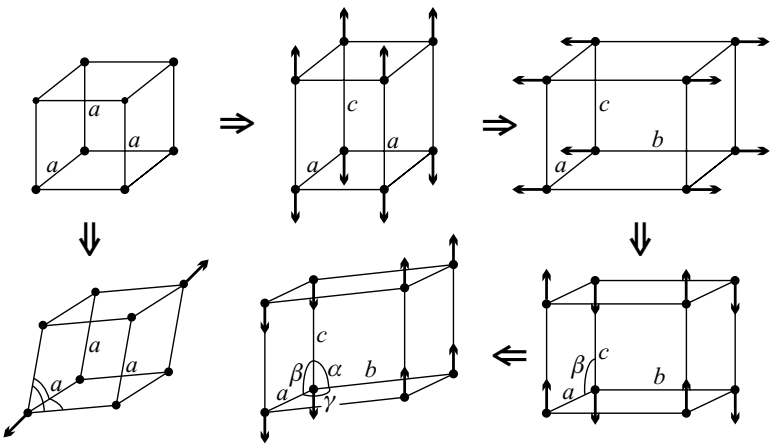


Fig. 5.25. Bodies of lower symmetry obtained through subsequent deformations of a cube. The direction of deformation forces are shown by the single arrows

Trigonal and hexagonal crystal systems are not yet included in the above scheme. By stretching the cubic lattice along a space diagonal, a rhombohedral lattice is obtained. Another small deformation leads to a lattice that possesses the symmetries of the monoclinic crystal system. Small deformations of a cubic (regular) lattice cannot lead to a hexagonal lattice. On the other hand, deformations of a hexagonal lattice may lead to a lattice showing orthorhombic symmetry – more precisely, to a base-centered orthorhombic Bravais lattice. Figure 5.26 gives a summary of this hierarchy of three-dimensional crystal systems.

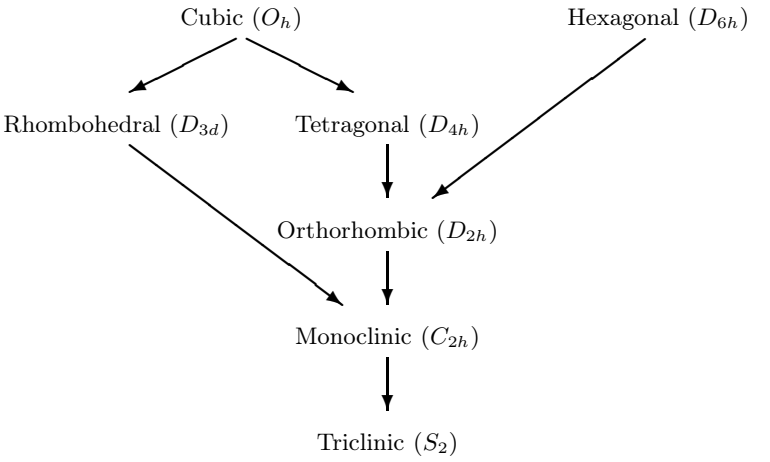


Fig. 5.26. Hierarchy of three-dimensional crystal systems

Each time the primitive cell is deformed, a symmetry is broken. Therefore the symmetries of a crystal system ranked lower in the hierarchy form a subgroup of the symmetries of the system ranked higher. The same hierarchy could have been derived by starting with the cubic and hexagonal point groups (O_h and D_{6h}), and choosing smaller and smaller subgroups. One must, however, exercise due care: although the symmetry group D_{3d} of the rhombohedral lattice is a subgroup of D_{6h}, the symmetry group of the hexagonal lattice, the former should not be considered to lie under the latter in the hierarchy as no small deformation of the hexagonal lattice can lead to a rhombohedral one.

In the above presentation of the hierarchy of crystal systems we started off with a simple cubic lattice and arrived at lower-symmetry lattices through subsequent deformations. Analogously, one can start with a body- or face-centered cubic lattice and track down deformed lattices with lower symmetries. It is a straightforward matter to prove that whichever cubic Bravais lattice type is taken, a suitably chosen deformation will take it into one of the tetragonal lattice types. Further deformations will transform any tetragonal lattice type into one of the orthorhombic lattices. The four orthorhombic lattice types are, in turn, deformed into either of the monoclinic lattice types. Finally, a small deformation of either monoclinic lattice type results in a triclinic lattice. Similarly to the case of the simple cubic lattice, when a face- or body-centered cubic lattice is stretched along the body diagonal, a rhombohedral lattice is obtained. As the hexagonal system has no centered type one can state that the hierarchy among crystal systems does not change when the various types within each system are taken into account. This hierarchy is of particular importance when the crystal classes are assigned to the crystal systems in the next section.

5.5 Full Symmetry of Crystals

Up to this point translation symmetries (elements of the group T_3) have been treated separately from the rotation and reflection symmetries (elements of the group G_0^3) of the Bravais lattice. Translations, rotations, and reflections all together make up the full symmetry group of a crystal's Bravais lattice. When a rotation or reflection (denoted by α) and a translation ($\boldsymbol{t}_n$) both take the lattice into itself then so does the succession of the two operations. In the Seitz notation[23] this composite symmetry operation is denoted by

$$\{\alpha|\boldsymbol{t}_n\}\,. \tag{5.5.1}$$

The order of the two operations is important as translations and rotations cannot usually be interchanged. Rotation (reflection) always comes first, and translation second. Because of this noncommutativity, the full symmetry group is the semidirect product of T_3 and G_0^3: $T_3 \rtimes G_0^3$. Analogously, for two-dimensional lattices the full symmetry group is $T_2 \rtimes G_0^2$.

When the translational symmetry of a crystal is identified and the primitive translation vectors are determined, the full crystal – lattice plus basis – has to be considered. However, when studying the rotation and reflection symmetries of Bravais lattices, empty lattices are considered – i.e., the basis is altogether ignored. When the basis is taken into account, some rotational symmetries may be broken: the crystal may not be invariant under each rotation that takes the underlying lattice into itself. Invariance is preserved only when finite but spherical atoms or precisely spherical molecules sit at the lattice points. Since this is generally not the case, depending on the shape of the group of atoms or molecules around each lattice point (i.e., the internal symmetry of the basis), certain rotation and reflection symmetries of the underlying lattice may be broken in the crystal. On the other hand, the crystal may have additional symmetries that do not exist in the Bravais lattice. That is why even after the separate discussion of translational and rotational symmetries new features appear when the full symmetry of the crystal is considered.

5.5.1 Screw Axes and Glide Planes

Once again, we shall first consider the two-dimensional case, in which patterns are repeated along two spatial directions. Such structures are shown in Fig. 5.27.

The absence of a mirror line is apparent at first sight: neither drawing is left-right symmetric. It is readily shown, however, that the pattern is brought into coincidence with itself when a reflection in the vertical midline is followed by a vertical translation. In addition to the well-known mirror line, a new

[23] F. Seitz, 1936.

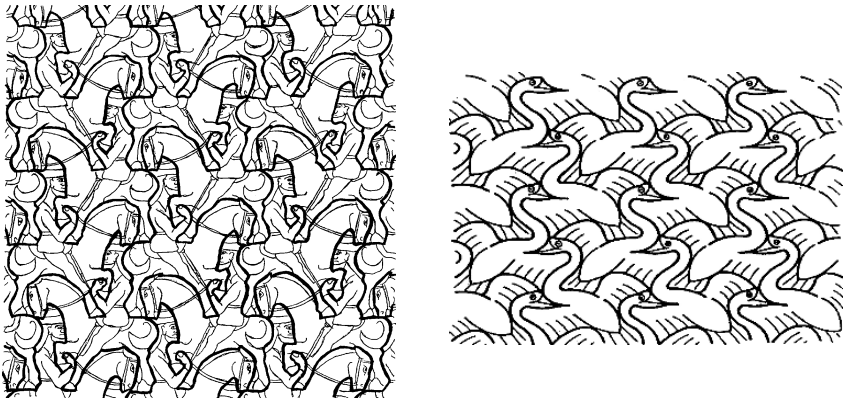


Fig. 5.27. Plane figures with glide lines on Escher's drawings

type of symmetry element appears in two-dimensional structures: the *glide line*, customarily denoted by g.

Glide reflection $\{m|\boldsymbol{v}\}$ is a composite symmetry operation in which a reflection m in a mirror line of the plane is followed by a translation $\boldsymbol{v}$ along the direction of the mirror line, where $\boldsymbol{v}$ is not an integer multiple of any primitive vector of the lattice. The succession of the two operations may be a symmetry transformation even though neither reflection m nor translation through $\boldsymbol{v}$ is a symmetry. Nevertheless, the length of the translation vector $\boldsymbol{v}$ cannot be arbitrary. As the resultant of two such operations is a pure translation, translation along the glide line must cover half the distance that characterizes periodicity along that direction.

In three-dimensional crystals two new symmetry operations may appear. The first one is *glide reflection* in space. This is a reflection in a plane followed by a translation parallel to the plane that is not an integral multiple of primitive translations.[24] The plane of the reflection is called the *glide plane*. Its Schoenflies symbol is σ^{g} and its international symbol is a, b, c, n, or d – depending on whether translation is along an edge, a face or body diagonal of the unit cell, through one-half or one-quarter of its length. Figure 5.28 shows a periodic structure with a glide plane.

In the other new symmetry operation α in $\{\alpha|\boldsymbol{v}\}$ is a rotation. A crystal structure is said to possess an *n-fold screw axis* if it is brought into coincidence with itself by an *n-fold screw rotation* (also called a *screw operation*) – i.e., rotation through $2\pi/n$ around some axis followed by a translation along the

[24] A reflection followed by a translation perpendicular to the mirror plane is not a new symmetry operation: it is equivalent to a reflection in a parallel plane. Nevertheless glide reflections are often chosen in such a way that $\boldsymbol{v}$ has both parallel and perpendicular components.

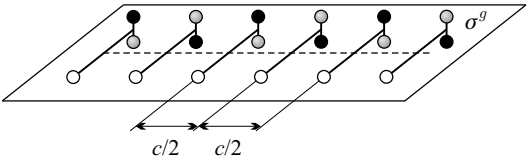


Fig. 5.28. Periodic structure with a glide plane

axis.[25] The symmetry element is denoted by C_n^{s} in the Schoenflies notation and by n_j in the international notation, where j runs over the values $1, 2, \ldots, n-1$, and refers to the succeeding translation – which is not arbitrary. The group property of symmetry operations implies that when the screw axis is twofold then rotation through 180° has to be followed by a translation along the axis through half the repeat distance (periodicity) in that direction. The symbol for this operation is 2_1. For a threefold screw axis translation has to be through either one-third or two-thirds of the repeat distance. The symbols for these operations are 3_1 and 3_2. In general, when the screw axis is n-fold, rotation through $2\pi/n$ is followed by a translation through j/n times the periodicity of the crystal. This is denoted by n_j. Figure 5.29 shows objects with a fourfold screw axis.

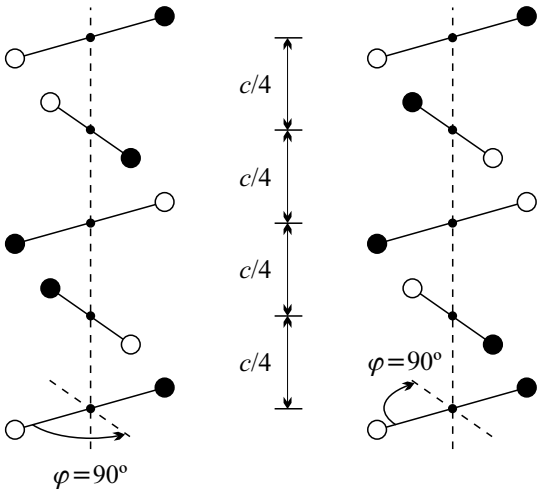


Fig. 5.29. Periodic structures with a fourfold screw axis. Rotation is clockwise in one case and counterclockwise in the other; both are followed by a translation of the crystal through one-quarter of the lattice constant

[25] A rotation followed by a translation perpendicular to the rotation axis is not a new symmetry operation, either: it is equivalent to a rotation around a parallel axis. Nevertheless screw operations are often chosen in such a way that $\boldsymbol{v}$ has both parallel and perpendicular components.

5.5.2 Point Groups of Crystals and Crystal Classes

Let us examine now how symmetries of the Bravais lattice are broken in the crystal when the basis has a lower symmetry than the Bravais lattice. As an example, we shall consider crystals whose underlying Bravais lattice is a simple cubic lattice and atoms at the lattice sites are surrounded by different atoms. Three different cases are shown in Fig. 5.30.

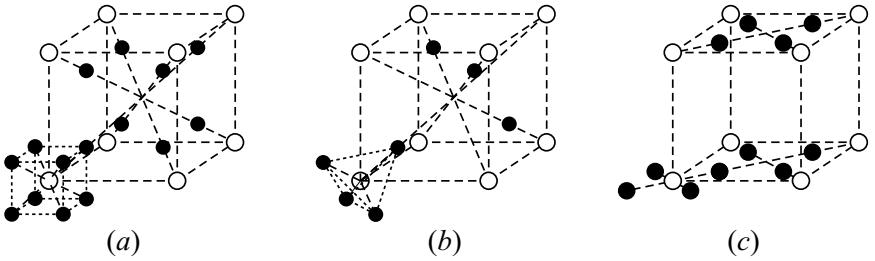


Fig. 5.30. Symmetry is either preserved or broken when a simple cubic Bravais cell is decorated with bases of various symmetries. (a) Symmetry of the point group O_h is preserved. (b) Symmetry is broken to T_d. (c) Symmetry is broken to D_{4h}

In the first case, among the atoms at the lattice points, atoms of another species ("black atoms") are placed in each of the eight equivalent points $\pm\xi, \pm\xi, \pm\xi$ of the space diagonals. Then the atom at the lattice points is surrounded by 8 black atoms in a cubic arrangement – therefore the basis and the crystal itself possess cubic symmetry (O_h). The same situation arises when 6 neighboring atoms are arranged along the edges, forming an octahedron.

However, when the atom at the lattice point is surrounded by only four black atoms then the symmetry of the basis can no longer be cubic – and so the cubic symmetry of the lattice cannot be entirely preserved. When the four atoms are arranged on the space diagonals in the manner shown in Fig. 5.30(b) – i.e., when the atom at the lattice point is surrounded by four black atoms arranged tetrahedrally – then only those rotational symmetries are preserved that take the tetrahedron into itself – namely, elements of the point group T_d. The atom at the lattice point is surrounded by four black atoms in part (c), too, however the black atoms are now placed on the diagonals of one face (the base). Threefold rotational symmetry around the space diagonal is thus broken – however, fourfold rotational symmetry around one of the edges reappears. The point group of the five-point basis is D_{4h}, which is also the point group of the rotational symmetries of the crystal.

These examples show that the crystal itself is not always invariant under each rotation and reflection of the point group of its Bravais lattice. Some symmetries may be broken because of the presence of a basis. The remaining operations form a group, which is necessarily a subgroup of the point group of

the Bravais lattice. Therefore this group has to be one of the crystallographic point groups listed in Table 5.5, as the subgroups of the seven Bravais groups are precisely the 32 point groups given there.

The same group may appear as the subgroup of more than one Bravais group. In such cases the group is considered to belong to the crystal system that is lowest in the hierarchy. This classification is based on the empirical finding that when a lattice is decorated with a basis whose symmetry is lower than that of the Bravais lattice, the lattice itself will be deformed, and the relations among the lattice parameters will correspond to the lower-ranked crystal system. For example, when the lattice points of a cubic crystal are decorated with molecules of tetragonal symmetry (as above), and thus the cubic symmetry is locally broken, the lattice itself will undergo a tetragonal deformation. More generally: if the equality of two quantities (in our case: lattice constants) is not required by symmetry then – apart from few exceptional, accidental cases – these quantities will be different.

Note that two-dimensional crystallographic point groups may also be classified into crystal systems based on their hierarchy. When a point group may appear as the subgroup of several Bravais groups then it is classified under the lowest one in the hierarchy. Thus point groups $6mm$, $3m$, 6, and 3 belong to the hexagonal system, $4mm$ and 4 to the square system, $2mm$ and m to the rectangular system, and 2 and 1 to the oblique system.

The knowledge of the rotation and reflection symmetries of a crystal is important because internal physical processes should take place in the same way in all directions that are equivalent due to symmetry – and so crystal symmetries manifest themselves in macroscopic properties, too. For example, the elasticity or conductivity tensor has to be invariant under the transformations that take the crystal into itself. We shall make extensive use of the severe restrictions imposed by this property on the tensor elements.

The symmetry operations which transform into each other those directions of the crystal that are equivalent in terms of macroscopic properties form the *point group of the crystal*. The elements of this group can be obtained by replacing screw rotations by simple rotations and glide reflections by mirror reflections. To see this it has to be born in mind that macroscopic properties depend only on the relative arrangement of the atoms along some direction. Thus two directions are equivalent even when they are transformed into one another by a screw rotation or glide reflection instead of a pure rotation or reflection.

Obviously, each element of the point group of the crystal must be an element of the point group (Bravais group) of the crystal's Bravais lattice. Hence the point group of the crystal must be a subgroup of the point group of the lattice. Thus the total number of such groups is the same as the number of subgroups of the Bravais groups of the seven crystal systems. Thirty-two subgroups are found – they correspond to the 32 crystallographic point groups – thus a crystal's point group is one of the crystallographic point groups, and the same notation is used for them.

Crystals with the same point group are said to belong to the same *crystal class*. The 32 crystal classes received their names from finite objects whose symmetry group is the respective point group. For example, the crystal class with symmetry O_h is called the *hexakisoctahedral class*.

5.5.3 Space Groups

Besides translations, rotations, and reflections, the full set of symmetries of a crystal contains screw operations and glide reflections as well. These operations form a group, provided group multiplication – as for the translation group and the point group – is understood as the succession of two operations. Simple geometrical considerations lead to

$$\{\alpha'|\boldsymbol{t}'_n\}\{\alpha|\boldsymbol{t}_n\} = \{\alpha'\alpha|\alpha'\boldsymbol{t}_n + \boldsymbol{t}'_n\}\,. \tag{5.5.2}$$

The inverse of a transformation is given by

$$\{\alpha|\boldsymbol{t}_n\}^{-1} = \{\alpha^{-1}|-\alpha^{-1}\boldsymbol{t}_n\}\,. \tag{5.5.3}$$

The group that contains all symmetry operations of a crystal is called its *space group*. Two-dimensional space groups are also called *plane groups* or *wallpaper groups*.

Plane Groups

To find all possible "space" groups (plane groups, wallpaper groups) of two-dimensional crystals, we shall construct planar patterns that are periodic in both directions. To this end, we shall start with a two-dimensional Bravais lattice and decorate it with bases of appropriate symmetry. The groups of such patterns are the G_2^2 groups. Some of them can be written as the semidirect product of the group T_2 of translations in the plane with a suitably chosen two-dimensional point group G_0^2. Suitable choice means that the point group in question must be a subgroup of the Bravais group of the Bravais lattice. This is possible when glide reflection is not a symmetry element of the plane group. The two upper drawings in Fig. 5.6 show that when an oblique cell is decorated with a figure that does not have a twofold axis – in other words, when the point group of the basis is lower than that of the Bravais lattice – then the crystal itself does not have a twofold axis, either. The plane group of such patterns is $p1$. On the other hand, when the basis has higher symmetry than the oblique Bravais lattice then the crystal is also invariant under the symmetry operations of the lattice. The plane group of such crystals (with a twofold rotation axis and a simple oblique Bravais lattice) is denoted by $p2$.

Two lattice types are possible for rectangular lattices (whose point group is $2mm$). Primitive and centered lattices can be equally decorated with a basis that has two perpendicular mirror lines. This leads to the space groups $p2mm$

and $c2mm$ (in short notation: pmm and cmm). However, the basis may have only one mirror line instead of two. In this case the plane groups are pm and cm. When a rectangular lattice is decorated with a basis that has no mirror lines then it is deformed spontaneously into an oblique lattice.

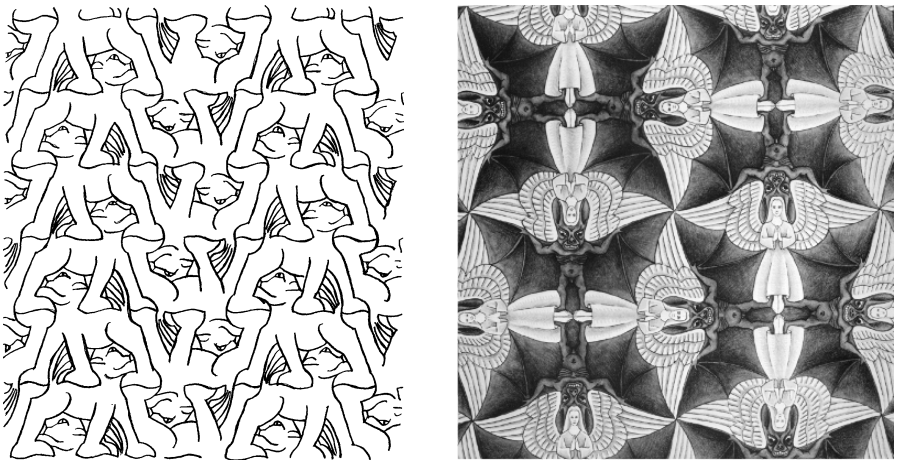
In square lattices, which are invariant under the rotations and reflections of the point group $4mm$, the basis itself may be invariant under $4mm$ operations. In this case the plane group of the crystal is $p4mm$ (in short notation: $p4m$). If the symmetry of the basis is lower, and only fourfold rotations are preserved then the plane group is $p4$. In principle, the symmetry of the basis could also be the subgroup m of the point group $4mm$. However, in this case the symmetry that links the two primitive vectors would be broken, and the square lattice would be deformed into a rectangular one. The resulting plane group would be pm.

When a hexagonal lattice is decorated with a basis that is also invariant under the symmetries of the point group $6mm$, then the plane group of the resulting crystal is $p6mm$ (in short notation: $p6m$). If the symmetry of the basis is lower, the lattice remains hexagonal as long as the point group of the basis is a subgroup of $6mm$: either $3m$, 6, or 3. The respective plane groups are $p3m$, $p6$, and $p3$. Motifs of symmetry $3m$ can be oriented in two different ways relative to the axes of the lattice, and so two plane groups can be generated. The threefold rotation axes are common in both of them, however the three mirror lines are either along or perpendicular to the three equivalent crystallographic axes. These are denoted by $p31m$ and $p3m1$.

Using translations, rotations, and reflections of the lattice, 13 two-dimensional space groups have been constructed above. They are called the *symmorphic groups* or *symmorphic plane groups*. There exist further, *nonsymmorphic* groups that contain glide reflections as well. When investigating the symmetries of a periodic motif one has to bear in mind that patterns of the symmorphic space group may also have glide lines although these do not pass through the group's characteristic n-fold axis. This is not the case for nonsymmorphic groups.

If a primitive rectangular lattice is decorated with such a basis that only one mirror line of the Bravais lattice is preserved and the other is replaced by a glide line, then the plane group is $p2mg$ (in short notation: pmg). It is equally possible that both mirror lines are replaced by glide lines, then the plane group is $p2gg$ (pgg). A motif with such symmetry is shown in Fig. 5.31(a). When one mirror line is completely absent and the other is replaced by a glide line, the pattern possesses the symmetries of the point group pg. Motifs with such symmetry were shown in Fig. 5.27.

A square lattice can be decorated with a basis in such a way that the mirror lines along the edges and diagonals of the square – intersecting each other at points of fourfold rotational symmetry – are no longer symmetries, but glide lines appear along the diagonals and parallel to the edges, while the



(*a*) (*b*)

Fig. 5.31. Two drawings of Escher with symmetries of two different plane crystallographic groups. (*a*) The plane group is $p2gg$ and the Bravais-lattice type is primitive rectangular for the figure with two glide lines. (*b*) The plane group is $p4gm$ and the Bravais-lattice type is square for the figure with one set of glide lines

fourfold rotation axes are preserved.[26] The plane group is then $p4gm$ ($p4g$). A pattern with such symmetry is presented in Fig. 5.31(*b*). Four nonsymmorphic groups are thus obtained, raising the total number of two-dimensional space groups to 17. Figure 5.32 shows periodic patterns that exhibit symmetries of two-dimensional space groups.

Three-Dimensional Space Groups

The group of symmetries of three-dimensional patterns that are periodic in three directions is the space group G_3^3. Once again, some of them may be written as the semidirect product of the group of translations T_3 with a three-dimensional point group G_0^3. This is possible when neither glide reflections nor screw rotations are among the symmetry elements. The group of rotations and reflections that take such a crystal into itself is either the point group of the Bravais lattice or one of its subgroups that belong to the same crystal system. By taking due care of the relative orientation of the rotational axes of the basis and the axes of the $6/mmm$ (D_{6h}) group in the hexagonal system, 73 different symmorphic space groups are constructed. Apart from the hexagonal system, the space groups are unambiguously specified by the Bravais-lattice type and the crystal class. This is reflected by their symbol in the international notation.

[26] Two perpendicular mirror lines are still present among the symmetry elements, however they go through the midpoints of adjacent perpendicular edges of the square, and their intersection points possess only twofold rotational symmetry.

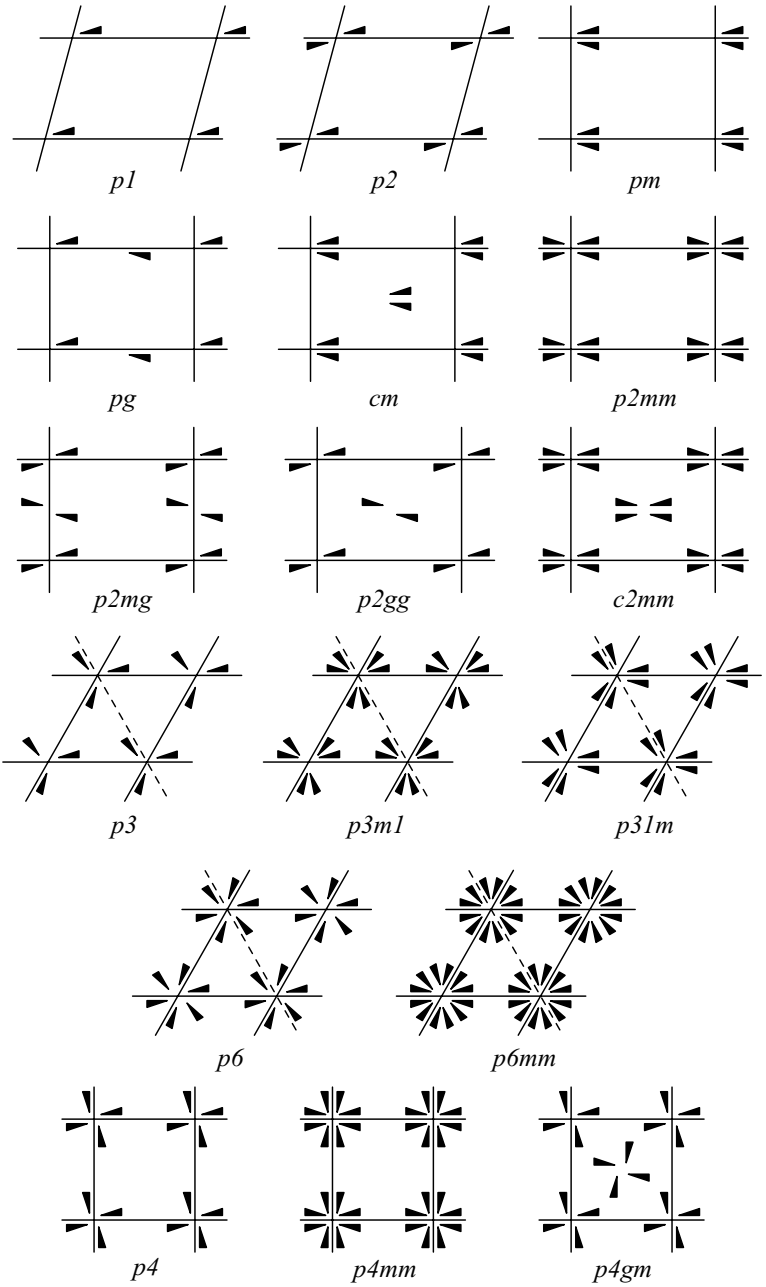


Fig. 5.32. Motifs showing the symmetries of plane groups

E.g., when a simple cubic lattice is decorated with a basis whose point group is $m\bar{3}m$, 432, $\bar{4}3m$, $m\bar{3}$, or 23, the space group of the crystal is $Pm\bar{3}m$, $P432$, $P\bar{4}3m$, $Pm\bar{3}$, or $P23$, respectively. When the original lattice type is body- or face-centered cubic, I or F appears in the symbol of the space group.

In addition to the 73 symmorphic space groups there exist 157 nonsymmorphic space groups, which contain screw operations and/or glide reflections. The total number of three-dimensional space groups is thus 230. This result was first worked out by the crystallographer J. S. Fedorov (1890) and independently of him by the mathematician A. M. Schoenflies (1891). We shall mercifully spare the reader a complete listing as it would lead too far afield. The 230 space groups and their symmetries are most extensively treated in Volume A of the series *International Tables for Crystallography*. As an example, we shall present the space groups that belong to class $m\bar{3}m$ (O_h) of the cubic crystal system.

The symmorphic space group $Pm\bar{3}m$ (in Schoenflies notation: O_h^1) is the space group of crystals whose underlying Bravais lattice is simple cubic and for which all elements of the point group $m\bar{3}m$ (O_h) – presented in Tables 5.1 and 5.4 – are symmetries. A nonsymmorphic space group is obtained if all rotations listed in Table 5.1 are symmetries but reflections listed in Table 5.4 take the crystal into itself only when followed by a translation along the space diagonal through one-half of its length. The generator of these rotoreflections is $\{I|\frac{1}{2}\frac{1}{2}\frac{1}{2}\}$. Compared to the space group $Pm\bar{3}m$, all rotations are pure rotations, however all mirror planes are replaced by glide planes. Consequently, its symbol is $Pn\bar{3}n$. In the Schoenflies notation space groups generated from the same point group are distinguished by a single superscript, its symbol is therefore O_h^2.

Different space groups are obtained when the space diagonals remain threefold rotation axes while the three principal axes and the six face diagonals become four- and twofold screw axes, respectively, with translation along the space diagonal. This means that the first 12 elements of Table 5.1 remain pure rotations while the next 12 are replaced by screw operations. If inversion remains a symmetry then along with it half of the reflections are also preserved and the other half are replaced by glide reflections. The resulting space group is called $Pm\bar{3}n$ (O_h^3). When, on the other hand, the symmetry under inversion is broken unless it is followed by a translation along the space diagonal, ordinary mirror planes of the space group $Pm\bar{3}n$ are replaced by glide planes – and vice versa. Consequently, the space group is denoted by $Pn\bar{3}m$ (O_h^4). There is no other way to obtain a cubic space group when starting with a simple lattice.

5.5.4 Symmetries of Magnetic Crystals

In the foregoing we have analyzed spatial transformations that take a point (x, y, z) of a crystal into an equivalent point. These form the group of spatial transformations. When ferroelectric or magnetic materials are studied, the

transformation of the electric and magnetic dipole moments need to be taken into account as well. As a simple example, A. V. SHUBNIKOV (1951) studied the symmetries of a system in which there is a two-valued variable $s = \pm 1$ at each site (x, y, z).

Because of the new variable, a new operation (R) appears, which is called *antisymmetry*. This takes the point characterized by the variables (x, y, z, s) into $(x, y, z, -s)$. Adding this operation to the customary symmetry operations leads to magnetic point groups and magnetic space groups or Shubnikov (space) groups. The two values ($s = \pm 1$) are often referred to as "black" and "white" – and then the group is called "black-and-white group".

Magnetic point groups contain rotations, reflections, and the antisymmetry operation. In addition to the 32 ordinary point groups of crystals (that do not contain the antisymmetry operation), another 32 appear which contain the antisymmetry operator R as a symmetry element. This is possible only when the values of s are both present in each lattice point. In magnetic terms this means that upward and downward magnetic moments are equally probable at each individual lattice point, and thus the average vanishes. This corresponds to the paramagnetic phase. Using black and white, these are the so-called gray groups. There are 58 other magnetic point groups that do not contain R itself only its combination with a rotation. They are the true magnetic groups in the sense that they may appear as points groups of ordered magnetic structures.

When translations are also taken into account, one may study magnetic (or black-and-white) lattices instead of ordinary Bravais-lattice types, in which all lattice points are equivalent. Besides displacements through translation vectors $\boldsymbol{t}_n$, operations of the form $\{R|\boldsymbol{v}\}$ are also allowed. Obviously, this may be a symmetry only when $\boldsymbol{v}$ is half a translation vector. In addition to the 14 ordinary Bravais-lattice types – in which lattice sites are of the same color –, 22 further Bravais-lattice types are found, which contain both black and white lattice sites. The new magnetic (or black-and white) lattice types with tetragonal and cubic structures are shown in Figs. 5.33 and 5.34.

Figure 5.33 shows that black-and-white Bravais lattices are built up of two interpenetrating Bravais lattices, a black and a white. In the case of a simple tetragonal lattice the relative displacement vector of the two sublattices is either one-half of the primitive vector along the fourfold axis, or the vector to

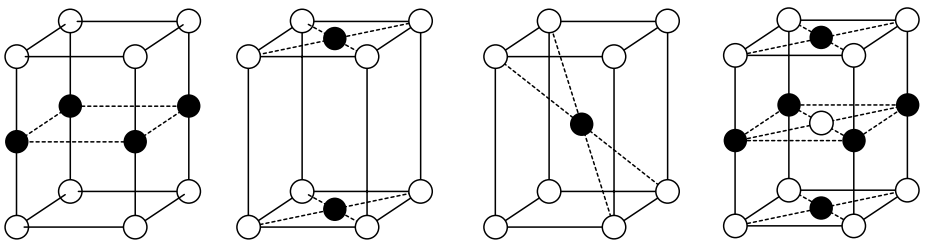


Fig. 5.33. Black-and-white Bravais lattices in the tetragonal crystal system

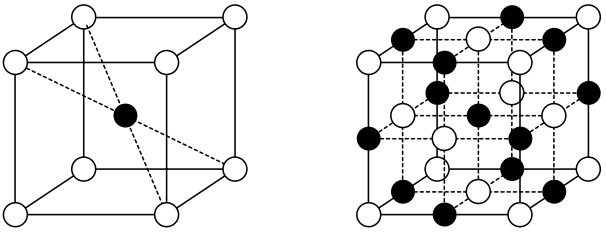


Fig. 5.34. Black-and-white Bravais lattices in the cubic crystal system

the center of the base, or the vector to the center of the primitive cell. In a body-centered tetragonal lattice there is just one possibility: the relative displacement vector of the two sublattices must be in the direction of the fourfold axis, and its magnitude must be one-half of the height. This is equivalent to a translation of the second sublattice through half the face diagonal.

In a simple cubic Bravais lattice the two interpenetrating sublattices are displaced by half the space diagonal. Each black atom is surrounded by eight white atoms, and vice versa. In a face-centered cubic lattice the relative displacement is the half of either edge vector. In this case each black atom is surrounded by six white ones.

The combination of magnetic Bravais lattices and magnetic point groups gives rise to 1651 magnetic (or black-and-white) space groups. Just like for point groups, 230 of them are ordinary space groups that do not contain the antisymmetry operation at all. There are the same number of paramagnetic (or gray) space groups, in which the antisymmetry operation is a symmetry element in itself. In the remaining 1191 space groups antisymmetry appears only in combination with rotations, reflections, or translations.

The possible symmetries of real magnetic systems are even more complicated than that. Black-and-white groups may, at most, purport to give the symmetry groups of magnetic systems that can be described by the Ising model, in which the magnetic moment is represented by a two-valued variable. In the general case the magnetic moment vector points in different directions on different magnetic atoms, giving rise to noncollinear magnetic structures. Moreover, account must be taken of the fact that magnetic moments are also transformed by rotations. It should also be borne in mind that the magnetic moment $\boldsymbol{m}$ is an axial vector, which does not transform as a true vector under reflections. (This last property is easily understood when the magnetic moment is considered to be produced by a current loop. Upon reflection, the component perpendicular to the mirror plane does not change sign but the component in the plane does.) All this leads to a great wealth of possibilities for symmetries in magnetic systems that cannot be listed here. The issue of possible magnetic structures will be briefly discussed in Chapter 14 on magnetically ordered systems.

Further Reading

1. C. J. Bradley and A. P. Cracknell, *The Mathematical Theory of Symmetry in Solids*, Clarendon Press, Oxford (1972).
2. *International Tables for Crystallography*, Volume A, *Space-Group Symmetry*, Edited by T. Hahn, Corrected reprint of the fifth edition, Springer-Verlag, Berlin (2005).
3. B. K. Vainshtein, *Modern Crystallography*, Vol. 1. *Fundamentals of Crystals. Symmetry and Methods of Structural Crystallography*, Second enlarged edition, Corrected 2nd printing, Springer-Verlag, Berlin (1996).
4. E. J. W. Whittaker, *Crystallography: An Introduction for Earth Science and Other Solid State Students*, Pergamon Press, Oxford (1981).

6

Consequences of Symmetries

In addition to the spatial symmetries discussed in the previous chapter a physical system may possess further symmetries, for example, it may be invariant under time reversal or gauge symmetries. While they might seem irrelevant to the structure of the crystal, they may nevertheless play an essential role in understanding the physical properties of the system. In general, the existence or lack of a symmetry may have fundamental influence on the system's behavior.

The connection between symmetries and physical properties is summarized by a fundamental postulate of crystal physics formulated back in the 19th century, known as Neumann's principle: *The symmetry elements of any observable physical property of a crystal must include all the symmetry elements of the point group of the crystal.* To put it otherwise: *The tensor representing any direction-dependent macroscopic physical property of a crystal must be invariant under the symmetry operations of the point group of the crystal.*

For physical quantities given in tensor form, such as resistivity, elastic constants, or optical properties, symmetry operations provide relations among the elements of the tensor. Using symmetry considerations alone it is possible to determine the number of different parameters that completely specify the tensor. For example, in the absence of a magnetic field, off-diagonal elements vanish and diagonal ones are identical in the two-index resistance tensor of a cubic crystal. Thus resistivity can be characterized by a single scalar quantity. In crystals with tetragonal symmetry, the resistivity tensor has two independent parameters. As we shall see, the four-index elasticity tensor is specified by three elastic constants in cubic crystals.

Moreover, taking symmetries into account facilitates the determination of the states of the system and provides selection rules for allowed processes. If the Hamiltonian of the system possesses some symmetry that is not observed in the ground state – i.e., if symmetry is spontaneously broken – then further conclusions can be drawn about the excited states of the system and the subgroup of preserved symmetries.

Using simple group theoretical considerations we shall first demonstrate that when symmetries are known, general statements can be made about the degeneracy of electronic states of the crystal, or about lifting this degeneracy. It will be shown that in the presence of an external perturbation, symmetries of the perturbation and of the system imply strict selection rules on the processes that may take place in the solid. Then we shall investigate the consequences of translational symmetry. Finally we shall discuss the implications of symmetry breaking.

6.1 Quantum Mechanical Eigenvalues and Symmetries

If a crystal possesses some symmetry its quantum mechanical Hamiltonian must possess the same symmetry. Knowledge of the symmetries often permits one to make general statements about the quantum mechanical eigenvalue problem without solving the Schrödinger equation explicitly. To this end one has to make recourse to group theory. We assume that the reader is familiar with the basics of group theory – nevertheless we shall summarize the most important theorems and relations for solid-state physics applications in Appendix D. Several theorems are concerned with the relation between the energy spectrum of a system and the irreducible representations of the group of its symmetry operations. Below, after recalling one of Wigner's theorems,[1] we shall show how they can be applied to the study of energy level splitting in solids and to the determination of selection rules.

6.1.1 Wigner's Theorem

Suppose that a system and its Hamiltonian $\mathcal{H}(\boldsymbol{r})$ are both invariant under the elements G_i of a symmetry group $\mathcal{G}$ – that is, operators $T(G_i)$ associated with the symmetry operations take the Hamiltonian into itself:

$$T(G_i)\mathcal{H}(\boldsymbol{r}) = \mathcal{H}(G_i^{-1}\boldsymbol{r}) = \mathcal{H}(\boldsymbol{r})\,. \tag{6.1.1}$$

This implies that the operators $T(G_i)$ commute with the Hamiltonian,

$$T(G_i)\mathcal{H}(\boldsymbol{r}) = \mathcal{H}(\boldsymbol{r})T(G_i)\,, \tag{6.1.2}$$

since by acting on an arbitrary function $\psi(\boldsymbol{r})$

$$\begin{aligned} T(G_i)\mathcal{H}(\boldsymbol{r})\psi(\boldsymbol{r}) &= \mathcal{H}(G_i^{-1}\boldsymbol{r})\psi(G_i^{-1}\boldsymbol{r}) \\ &= \mathcal{H}(\boldsymbol{r})\psi(G_i^{-1}\boldsymbol{r}) \\ &= \mathcal{H}(\boldsymbol{r})T(G_i)\psi(\boldsymbol{r})\,. \end{aligned} \tag{6.1.3}$$

Let $\psi(\boldsymbol{r})$ be an eigenfunction of the Hamiltonian of the system with eigenvalue ε, i.e.,

[1] E. P. Wigner, 1927.

$$\mathcal{H}(\boldsymbol{r})\psi(\boldsymbol{r}) = \varepsilon\psi(\boldsymbol{r})\,. \tag{6.1.4}$$

Operating $T(G_i)$ on both sides and using (6.1.3) gives

$$T(G_i)\mathcal{H}(\boldsymbol{r})\psi(\boldsymbol{r}) = \varepsilon T(G_i)\psi(\boldsymbol{r}) \tag{6.1.5}$$

and

$$\mathcal{H}(\boldsymbol{r})T(G_i)\psi(\boldsymbol{r}) = \varepsilon T(G_i)\psi(\boldsymbol{r})\,, \tag{6.1.6}$$

that is, $T(G_i)\psi(\boldsymbol{r})$ is also an eigenfunction, with the same eigenvalue. This result is summarized by Wigner's theorem: *If $\psi(\boldsymbol{r})$ is an eigenfunction of the Hamiltonian with an energy eigenvalue ε then for any symmetry operation $T(G_i)$ that leaves the Hamiltonian invariant $T(G_i)\psi(\boldsymbol{r}) = \psi(G_i^{-1}\boldsymbol{r})$ is also an eigenfunction, with the same energy.*

By operating all elements of the symmetry group $\mathcal{G}$ on the wavefunction $\psi(\boldsymbol{r})$, a set of eigenfunctions with the same eigenvalue is obtained. One may then construct a representation of the group on the space of these functions. If the representation is irreducible then the degeneracy of the energy levels is not accidental but the consequence of symmetry. When, on the other hand, the representation is not irreducible, then after its reduction to irreducible representations symmetry operations will mix only those wavefunctions that belong to the same representation, and only their eigenvalues will be necessarily equal. Higher degeneracy may appear only accidentally. It is plausible that by choosing the potential (interaction term) in the Hamiltonian in the most general form consistent with symmetry accidental degeneracy of the energy levels is lifted, and only the degeneracy imposed by symmetries survives. In such cases the representation of the symmetry group is irreducible on the wavefunctions of the selected energy value.

Consequently, energy levels can be characterized by the irreducible representation according to which the wavefunctions transform, and the degree of degeneracy is equal to the dimension of the irreducible representation associated with the energy level. This also means that when the symmetries of the system and the irreducible representations of the symmetry group are known, one can determine the degree of degeneracy of the energy levels, as well as the way they are split in the presence of a lower-symmetry external perturbing field.

6.1.2 Splitting of Atomic Levels in Crystals

The state of atomic electrons in a spherical potential is characterized by the principal quantum number $n = 1, 2, 3, \ldots$, the azimuthal quantum number $l = 0, 1, 2, \ldots$ (that specifies the orbital angular momentum of the states s, p, $d, \ldots$), the magnetic quantum number $m_l = -l, -l+1, \ldots, l-1, l$, and the spin quantum number $m_s = \pm 1/2$. To determine energy levels, in addition to the nucleus one must also take into account the effects of the other electrons. Approximating the latter by a spherically symmetric effective

potential, the potential field is invariant under all rotations around any axis through the origin, irrespective of the angle. The group of these symmetry operations is equivalent to the special orthogonal group SO(3). Its irreducible representations can be characterized by the eigenvalues of the orbital angular momentum operator L. As the dimension of the irreducible representation associated with quantum number l is $2l+1$, the corresponding energy level is $2l+1$-fold degenerate. Owing to the radial part of the wavefunction, energy obviously depends on the principal quantum number n as well, however, it is independent of m_l.

Instead of a single isolated atom consider one at a lattice site of a regular crystal, where the effects of its neighbors may no longer be neglected. Even when electrons are localized on atoms and effects of overlapping between electronic states can be ignored, the electrostatic field produced by neighboring atoms, the *crystal field* may change the electronic states. This potential possesses the local symmetry of the site in question, that is, the symmetry in the arrangement of the neighboring atoms. This is usually lower than perfect spherical symmetry, therefore, within a crystal, the free ion's originally $2l+1$-fold degenerate energy level of quantum number l is split into several sublevels. This phenomenon is called *crystal-field splitting*.

The amount of splitting depends on the strength of the potential at the atom, nevertheless the character of splitting – i.e., the number and degeneracy of the sublevels into which the atomic level is split – are determined solely by the symmetries of the atomic position. Without knowing the exact form of the crystal field, these can be specified from the symmetry properties of the crystal. This is of particular importance as it allows for the determination of atomic positions.

As an example, consider an atom in a cubic environment for which the group of local symmetries is O_h. This is the case for an atom at the center of the Bravais cell in a face-centered cubic crystal. As we shall see in the next chapter, this position is called an octahedral site. Then the crystal field $V_{\text{cr field}}(r,\theta,\varphi)$ acting on the electrons orbiting the atom possesses cubic symmetry, i.e., it is invariant under the rotations and reflections listed in Tables 5.1 and 5.4. The simplest such potential contains the lowest-order polynomial that shows cubic symmetry but is not invariant under an arbitrary rotation:

$$\begin{aligned} V_{\text{cr field}} &= A(r) + D(x^4 + y^4 + z^4 - \frac{3}{5}r^4) + \dots \\ &= A(r) + Dr^4\Big[1 - \frac{5}{4}\sin^4\theta\sin^2 2\varphi - \frac{5}{4}\sin^2 2\theta\Big]. \end{aligned} \tag{6.1.7}$$

The solutions to the Schrödinger equation with this potential must transform according to the irreducible representations of the symmetry group of the cubic crystal. If the $2l+1$ states of a free atom with angular momentum l are used as a set of basis functions, the representation of the symmetry group of the cubic crystal (which contains a finite number of rotations through 90°,

120°, and 180°) is usually reducible on it. Thus the splitting of the state with orbital angular momentum l depends on how this reducible representation is reduced to the sum of the irreducible representations of the cubic symmetry group. The remaining degeneracy is determined by the dimensions of these irreducible representations. This reduction is readily calculated using the characters of the representations.

According to the formulas given in Appendix C, the character of a rotation through angle φ is

$$\chi_l(\varphi) = \sum_{m=-l}^{l} \mathrm{e}^{\mathrm{i}m\varphi} = \frac{\sin(l+\frac{1}{2})\varphi}{\sin\frac{1}{2}\varphi} \tag{6.1.8}$$

in the representation of index l of the rotation group.

Taking this expression for integer values of l at angles $\varphi = 0$, π, $\pi/2$, and $2\pi/3$, which appear among the symmetries of the cubic system,

$$\chi_l(0) = 2l+1\,, \tag{6.1.9-a}$$

$$\chi_l(\pi) = (-1)^l\,, \tag{6.1.9-b}$$

$$\chi_l(\pi/2) = \begin{cases} (-1)^{l/2} & \text{for } l \text{ even}\,, \\ (-1)^{(l-1)/2} & \text{for } l \text{ odd}\,, \end{cases} \tag{6.1.9-c}$$

$$\chi_l(2\pi/3) = \begin{cases} 1 & \text{for } l = 3k\,, \\ 0 & \text{for } l = 3k+1\,, \\ -1 & \text{for } l = 3k+2\,. \end{cases} \tag{6.1.9-d}$$

The characters for the 48 elements of the cubic symmetry group O_h are given in Table 6.1. When calculating them one must take into account that the elements of the cubic group can be divided into ten classes (see Appendix D). The first class contains the identity element; the second – rotations through 180° around the three equivalent directions $\langle 100\rangle$ (C_{4m}^2, $m = x, y, z$); the third – rotations through 90° and 270° around the directions $\langle 100\rangle$ (C_{4m} and C_{4m}^3); the fourth – rotations through 180° around the six equivalent directions $\langle 110\rangle$ (C_{2p}, $p = a, b, c, d, e, f$); the fifth – rotations through 120° and 240° around the directions $\langle 111\rangle$ (C_{3j} and C_{3j}^2, $j = a, b, c, d$). The characters for these classes can be read off from the formulas in (6.1.9). Elements of the other five groups are obtained by multiplying the above by the inversion operation. A basis function is odd (even) under inversion when the azimuthal quantum number is odd (even).

To proceed with the reduction process one must also know the characters of the irreducible representations of the group O_h. These are listed in Table D.1, in Appendix D. The general rules of reduction are also given there. Some straightforward algebra leads to the results in Table 6.2.

Table 6.1. Character table for the representation of the point group O_h using the spherical harmonics $Y_l^m(\theta,\varphi)$ as basis functions. Only the first few values of the quantum number l are listed

l	E	$3C_{4m}^2$	$6C_{4m}$	$6C_{2p}$	$8C_{3j}$	I	$3\sigma_m$	$6S_{4m}$	$6\sigma_p$	$8S_{3j}$
0	1	1	1	1	1	1	1	1	1	1
1	3	−1	1	−1	0	−3	1	−1	1	0
2	5	1	−1	1	−1	5	1	−1	1	−1
3	7	−1	−1	−1	1	−7	1	1	1	−1
4	9	1	1	1	0	9	1	1	1	0
5	11	−1	1	−1	−1	−11	1	−1	1	1
6	13	1	−1	1	1	13	1	−1	1	1

Table 6.2. Reduction of representation D_l of the rotation group to the irreducible representations of the group O_h. Only the first few values of the quantum number l are listed

l	Reduction	Dimension
0	$D_0 = \Gamma_1$	1
1	$D_1 = \Gamma_{15}$	3
2	$D_2 = \Gamma_{12} \oplus \Gamma'_{25}$	$5 = 2 + 3$
3	$D_3 = \Gamma'_2 \oplus \Gamma_{15} \oplus \Gamma_{25}$	$7 = 1 + 3 + 3$
4	$D_4 = \Gamma_1 \oplus \Gamma_{12} \oplus \Gamma'_{15} \oplus \Gamma'_{25}$	$9 = 1 + 2 + 3 + 3$

As an application, consider an atom with a 3d-electron placed in a cubic environment, and examine the possible states of the electron. According to Hund's first and second rules, $S = 1/2$, $L = 2$ in ground state – thus the state is ^{2}D. In a free atom this tenfold degenerate state is split by the spin–orbit interaction into a fourfold degenerate $J = 3/2$ state and a sixfold degenerate $J = 5/2$ state. According to Hund's third rule, the state with angular momentum $J = 3/2$ is the ground state. However, this is not necessarily the case for atoms in a crystal. The reason for this is that in transition metals – especially the elements of the iron group – with partially filled 3d, 4d, or 5d shells the crystal-field splitting for d electron levels (that are considered to belong to the core) is on the order of $\Delta E_{\text{cr field}} \sim 10^3\,\text{cm}^{-1} \sim 0.1\,\text{eV}$, which is larger than the energy correction due to spin–orbit interaction, $\Delta E_{\text{s–o}} \sim 10^2\,\text{cm}^{-1} \sim 0.01\,\text{eV}$. Thus to determine the state of the electron, the effects of crystal field in the configuration specified by Hund's first and second rules have to be taken into account first, and spin–orbit interactions second.

Writing the spatial part of the wavefunction of the electrons of azimuthal quantum number l in a free atom as

$$\psi_{nlm_l}(r,\theta,\varphi) = R_{nl}(r)Y_l^{m_l}(\theta,\varphi) = R_{nl}(r)P_l^{m_l}(\cos\theta)\mathrm{e}^{\mathrm{i}m_l\varphi}\,, \qquad (6.1.10)$$

the associated Legendre polynomials for the d-state ($l = 2$) are

$$P_2^{m_l}(\cos\theta) = \begin{cases} \frac{1}{3}\sin^2\theta & m_l = \pm 2, \\ \frac{1}{3}\sin\theta\cos\theta & m_l = \pm 1, \\ \frac{1}{2}(3\cos^2\theta - 1) & m_l = 0. \end{cases} \tag{6.1.11}$$

These functions transform according to the representation D_2 of the rotation group. In a cubic environment, this fivefold (or, when spin is also taken into account, tenfold) degenerate level is split into two levels that transform according to Γ_{12} and Γ'_{25}. A two- and a threefold (or, with spin, a four- and a sixfold) degenerate energy level are thus obtained.

It is readily shown that the basis functions of the irreducible representations are the symmetric and antisymmetric combinations of the functions with $m_l = \pm 2$ and $m_l = \pm 1$, plus the function with $m_l = 0$:

$$\begin{aligned}
\psi_{xy} &= R_{n2}(r)\frac{-\mathrm{i}}{\sqrt{2}}\left(Y_2^2 - Y_2^{-2}\right) = \left(\frac{15}{16\pi}\right)^{1/2} R_{n2}(r)\sin^2\theta\sin 2\varphi \\
&= \left(\frac{15}{4\pi}\right)^{1/2} R_{n2}(r)xy/r^2\,, \\
\psi_{x^2-y^2} &= R_{n2}(r)\frac{1}{\sqrt{2}}\left(Y_2^2 + Y_2^{-2}\right) = \left(\frac{15}{16\pi}\right)^{1/2} R_{n2}(r)\sin^2\theta\cos 2\varphi \\
&= \left(\frac{15}{16\pi}\right)^{1/2} R_{n2}(r)(x^2 - y^2)/r^2\,, \\
\psi_{xz} &= R_{n2}(r)\frac{-1}{\sqrt{2}}\left(Y_2^1 - Y_2^{-1}\right) = \left(\frac{15}{4\pi}\right)^{1/2} R_{n2}(r)\sin\theta\cos\theta\cos\varphi \\
&= \left(\frac{15}{4\pi}\right)^{1/2} R_{n2}(r)xz/r^2\,, \\
\psi_{yz} &= R_{n2}(r)\frac{\mathrm{i}}{\sqrt{2}}\left(Y_2^1 + Y_2^{-1}\right) = \left(\frac{15}{4\pi}\right)^{1/2} R_{n2}(r)\sin\theta\cos\theta\sin\varphi \\
&= \left(\frac{15}{4\pi}\right)^{1/2} R_{n2}(r)yz/r^2\,, \\
\psi_{z^2} &= R_{n2}(r)Y_2^0(\theta,\varphi) = \left(\frac{5}{16\pi}\right)^{1/2} R_{n2}(r)(3\cos^2\theta - 1) \\
&= \left(\frac{5}{16\pi}\right)^{1/2} R_{n2}(r)(3z^2 - r^2)/r^2\,.
\end{aligned} \tag{6.1.12}$$

These wavefunctions are real. The corresponding charge distributions are schematically shown in Fig. 6.1.

Wavefunctions ψ_{xy}, ψ_{yz}, and ψ_{xz} are equivalent in a cubic environment and transform according to the irreducible representation Γ'_{25} of the cubic

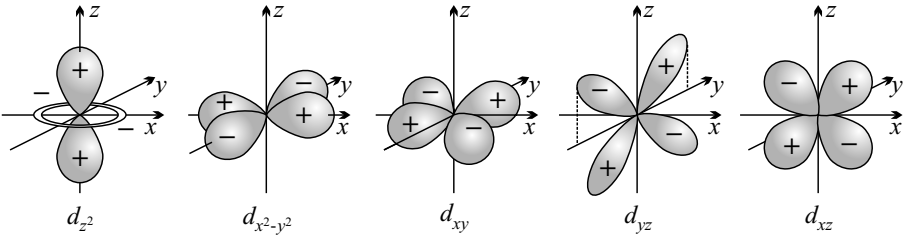


Fig. 6.1. Schematic representation of d-state wavefunctions

point group. In chemical nomenclature, this representation is also denoted by $T_{2\mathrm{g}}$ – which is why these states are often called $t_{2\mathrm{g}}$-states.

Functions $\psi_{x^2-y^2}$ and ψ_{z^2} transform according to the irreducible representation Γ_{12}, and they, too, have the same energy in a cubic environment. Chemical nomenclature uses the symbol E_{g} for this representation and the name e_{g}-state for the corresponding states.

As $t_{2\mathrm{g}}$- and e_{g}-states belong to different irreducible representations, their energies are usually different. Which of them is lower depends on the spatial arrangement of the surrounding atoms. When nearest neighbors are located along crystallographic axes forming an octahedral environment, $t_{2\mathrm{g}}$-states tend to be of lower energy, as wavefunctions overlap to a smaller extent and the Coulomb repulsion is weaker. The pattern of energy-level splitting is similar when a single 3d-electron is missing – however, the order of the levels is then reversed.

Using the wavefunctions given in (6.1.12), the expectation value for the z component of the angular momentum contains either the integral

$$\int_0^{2\pi} \sin(m_l\varphi)\,\frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial\varphi}\sin(m_l\varphi)\,\mathrm{d}\varphi \tag{6.1.13}$$

or

$$\int_0^{2\pi} \cos(m_l\varphi)\,\frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial\varphi}\cos(m_l\varphi)\,\mathrm{d}\varphi\,. \tag{6.1.14}$$

Both of them vanish, therefore

$$\langle L_z\rangle = 0\,. \tag{6.1.15}$$

The vanishing of the two other components can be shown along similar lines. The orbital angular momentum is thus said to be *quenched*. This explains the curiosity presented in Chapter 3 stating that one should ignore orbital angular momentum contributions to the effective magneton number of transition-metal ions. The hitherto neglected spin–orbit interaction may slightly polarize states in a crystal field, giving rise to a slight change in the effective magneton number.

6.1.3 Spin Contributions to Splitting

Rare earth metals with a partially filled 4f shell – lanthanoids (lanthanides)[2] – show a markedly different behavior from transition metals. Because of the screening effect of electrons on the 5s and 5p shells, splitting caused by the crystal field due to neighboring atoms is typically on the order of just $\Delta E_{\text{cr field}} \sim 10^2\,\text{cm}^{-1} \sim 0.01\,\text{eV}$, whereas splitting caused by the spin–orbit interaction is an order of magnitude larger, $\Delta E_{\text{s–o}} \sim 10^3\,\text{cm}^{-1} \sim 0.1\,\text{eV}$. For lanthanoids – as well as actinoids – the effect of the crystal field is thus smaller than that of the spin–orbit interaction; therefore the ground-state electron configuration has to be specified using Hund's three rules, and then the effects of the crystal field have to be treated as a perturbation, with J fixed.

For example, when there is a single electron on the 4f shell, then, according to Hund's rules the quantum numbers of the ground state are $S = 1/2$, $L = 3$, $J = L - S = 5/2$ ($^2\text{F}_{5/2}$ state), and the state with the same L and S but $J = 7/2$ is the first excited state. To determine the crystal-field splitting of the sixfold degenerate ground state, the behavior of the wavefunction under rotations should be known not only for integer values of the angular momentum but also half-integer values, which will be denoted by j. Relations (F.1.29) and (6.1.8) are still valid:

$$\chi_j(\varphi) = \frac{\sin(j+\frac{1}{2})\varphi}{\sin\frac{1}{2}\varphi}\,. \tag{6.1.16}$$

This implies

$$\chi_j(\varphi + 2\pi) = -\chi_j(\varphi)\,, \tag{6.1.17}$$

that is, because of the spinor character of spin, rotation through 2π is not an identity transformation, only rotation through 4π is. Thus, in the spirit of the discussion in Appendix D, the group of symmetries has to be extended to the double group in which rotation through 2π (denoted by $\overline{E}$) and its products with ordinary rotations also appear.

The characters for rotations compatible with cubic symmetry are

$$\chi_j(0) = 2j + 1\,, \tag{6.1.18-a}$$

$$\chi_j(\pi) = 0\,, \tag{6.1.18-b}$$

[2] Instead of the commonly used names *lanthanides* and *actinides*, the International Union of Pure and Applied Chemistry (IUPAC) recommends the usage of *lanthanoids* and *actinoids* for the groups of atoms from La to Lu and from Ac to Lr.

$$\chi_j(\pi/2) = \begin{cases} \sqrt{2} & \text{for } j = 1/2, 9/2, \dots, \\ 0 & \text{for } j = 3/2, 7/2, 11/2, \dots, \\ -\sqrt{2} & \text{for } j = 5/2, 13/2, \dots, \end{cases} \qquad (6.1.18\text{-c})$$

$$\chi_j(2\pi/3) = \begin{cases} 1 & \text{for } j = 1/2, 7/2, \dots, \\ -1 & \text{for } j = 3/2, 9/2, \dots, \\ 0 & \text{for } j = 5/2, 11/2, \dots. \end{cases} \qquad (6.1.18\text{-d})$$

From these relations, one may determine the characters for the elements of the double group O' of the octahedral group for various values of j. These are listed in Table 6.3.

Table 6.3. Characters for the elements of the double group O' on the basis of wavefunctions with half-integer angular momentum, for various values of j

j	E	$\overline{E}$	$3C_{4m}^2 + 3\overline{C}_{4m}^2$	$6C_{4m}$	$6\overline{C}_{4m}$	$6C_{2p} + 6\overline{C}_{2p}$	$8C_{3j}$	$8\overline{C}_{3j}$
1/2	2	−2	0	$\sqrt{2}$	$-\sqrt{2}$	0	1	−1
3/2	4	−4	0	0	0	0	−1	1
5/2	6	−6	0	$-\sqrt{2}$	$\sqrt{2}$	0	0	0
7/2	8	−8	0	0	0	0	1	−1
9/2	10	−10	0	$\sqrt{2}$	$-\sqrt{2}$	0	−1	1

Using the character table of the irreducible representations of the double group O' – Table D.2 of Appendix D –, the reductions listed in Table 6.4 are readily established.

Turning back to the case when the f ($l = 3$) shell contains a single electron, the total angular momentum may be either $j = 5/2$ or $j = 7/2$. The wavefunctions of the $j = 5/2$ level that transform according to the representation $D_{5/2}$ will, in a cubic crystal field, go over into functions that transform

Table 6.4. Reduction of the representations D_j with half-integer j of the continuous rotation group to the irreducible representations of the double group O'

j	Reduction	Dimension
1/2	$D_{1/2} = \Gamma_6$	2
3/2	$D_{3/2} = \Gamma_8$	4
5/2	$D_{5/2} = \Gamma_7 \oplus \Gamma_8$	$6 = 2 + 4$
7/2	$D_{7/2} = \Gamma_6 \oplus \Gamma_7 \oplus \Gamma_8$	$8 = 2 + 2 + 4$
9/2	$D_{9/2} = \Gamma_6 \oplus \Gamma_8 \oplus \Gamma_8$	$10 = 2 + 4 + 4$

according to the representations Γ_7 and Γ_8, therefore the energy level is split into a twofold and a fourfold degenerate level. Similarly, the $j = 7/2$ level is split into two twofold (Γ_6 and Γ_7) and a fourfold degenerate (Γ_8) level. The full spectrum thus contains three twofold and two fourfold degenerate levels. The resulting level structure is shown in Fig. 6.2.

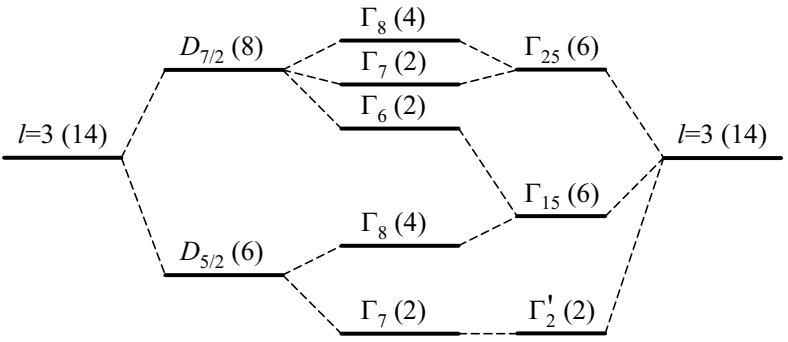


Fig. 6.2. Splitting of the 14-fold degenerate energy level of a single electron on the f shell in a cubic crystal field. On the left-hand side, spin–orbit interaction is taken into account first, and crystal-field splitting second; this order is reversed on the right-hand side. Numbers in parentheses show the degree of degeneracy

The right part of the diagram shows another approach: the 14-fold degenerate state of quantum numbers $S = 1/2$ and $L = 3$, which initially transforms according to the representation D_3, is first split by the crystal field, and then the spin–orbit interaction is also turned on. In a cubic crystal field the representation D_3 is decomposed into the sum of the representations Γ_2', Γ_{15}, and Γ_{25} – which correspond to a twofold and two sixfold degenerate levels (when spins are also counted). To account for further splitting caused by the spin–orbit interaction, spin is associated with a wavefunction that transforms according to $D_{1/2}$. To reduce the product of the spatial and spin wavefunctions, one has to exploit the relations

$$\begin{aligned} \Gamma_2' \otimes D_{1/2} &= \Gamma_7 \,, \\ \Gamma_{15} \otimes D_{1/2} &= \Gamma_6 \oplus \Gamma_8 \,, \\ \Gamma_{25} \otimes D_{1/2} &= \Gamma_7 \oplus \Gamma_8 \,. \end{aligned} \tag{6.1.19}$$

This leads to the same results as above, where the effects of the two interactions were taken into account in reversed order. The reason for this is that group-theoretical considerations say nothing about the order of split levels. The choice of one approach or the other is dictated by the relative strengths of the interactions: the stronger is taken into account first, and then the weaker is treated as a perturbation.

As we have seen, when an ion with a single f-electron is placed into a lattice with cubic symmetry, crystal-field splitting reduces the sixfold degeneracy of the ground state to twofold. Because of this, at low temperatures the

magnetic properties of the atom are similar to those of a spin-1/2 particle. At higher temperatures, where thermal energies are comparable with the energy splitting, each of the six $J = 5/2$ states are excited thermally. The temperature dependence of susceptibility therefore deviates from the simple Curie law.

6.1.4 Kramers' Theorem

The previous calculation of the possible states of a single electron showed that whatever kind of splitting is studied and however low the symmetry of the local environment, if spin is also taken into account then even spin–orbit interaction fails to split atomic energy levels completely – and the degree of degeneracy will be an even number for each state. This finding is generalized by Kramers' theorem:[3] *In the absence of a magnetic field – which could break time-reversal symmetry – each energy level is at least doubly degenerate in any system that contains an odd number of electrons.*

To verify this statement it should be recalled that the quantum mechanical operator of time reversal T is antiunitary, that is, it can be written as the product of a unitary operator U and the operator K_0 of complex conjugation.[4] For spinless particles T effectively transforms the wavefunction into its complex conjugate:

$$T\psi(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, t) = \psi^*(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, -t)\,. \tag{6.1.20}$$

For spin-1/2 particles the components of the wavefunction represented by a two-component spinor become mixed as time reversal changes the sign of spin:

$$T\boldsymbol{s}T^{-1} = -\boldsymbol{s}\,. \tag{6.1.21}$$

Following the customary quantization conventions, the eigenvalues of the spin z component are chosen as spin states:

$$K_0 s_x K_0 = s_x, \qquad K_0 s_y K_0 = -s_y, \qquad K_0 s_z K_0 = s_z\,. \tag{6.1.22}$$

Then the unitary part U of the time-reversal operator has to rotate the spin through π around the y-axis. One possible choice is

$$T = \mathrm{i}\mathrm{e}^{-\mathrm{i}\pi s_y} K_0 = \sigma_y K_0\,, \tag{6.1.23}$$

where σ_y is the corresponding Pauli matrix. When time reversal is performed on the eigenstates of the spin z components, the spin component changes its sign and the wavefunction receives a further spin-dependent factor. A convenient choice is

[3] H. A. Kramers, 1930.

[4] For antiunitary operators $\big(T\phi, T\psi\big) = \big(\psi, \phi\big)$, where $(,)$ denotes the scalar product, and $T(a\phi + b\psi) = a^* T\phi + b^* T\psi$.

$$T\psi(\boldsymbol{r}, s_z, t) = (-\mathrm{i})^{2s_z}\psi^*(\boldsymbol{r}, -s_z, -t)\,. \tag{6.1.24}$$

These expressions can be generalized to several particles. Either the form

$$T = \sigma_{1y}\sigma_{2y}\ldots K_0 \tag{6.1.25}$$

is used, or, when the particles are in some eigenstates of the spin z component,

$$T\psi(\boldsymbol{r}_1, s_{1z}, \boldsymbol{r}_2, s_{2z}, \ldots, t) = (-\mathrm{i})^{2(s_{1z}+s_{2z}+\ldots)}\psi^*(\boldsymbol{r}_1, -s_{1z}, \boldsymbol{r}_2, -s_{2z}, \ldots, -t)\,. \tag{6.1.26}$$

This implies that for the eigenstates of a system with n electrons,

$$T^2\psi = (-1)^n\psi\,. \tag{6.1.27}$$

As T contains a rotation of the spin by π, T^2 corresponds to a 2π rotation. However, a spinor is not transformed into itself but into its negative by such a rotation. This explains the factor $(-1)^n$.

Time reversal transforms a Hamiltonian that contains position, momentum, and spin variables $\boldsymbol{r}_i$, $\boldsymbol{p}_i$, and $\boldsymbol{s}_i$ into

$$T\mathcal{H}(\boldsymbol{r}_i, \boldsymbol{p}_i, \boldsymbol{s}_i)T^{-1} = \mathcal{H}(\boldsymbol{r}_i, -\boldsymbol{p}_i, -\boldsymbol{s}_i)\,. \tag{6.1.28}$$

In the absence of a magnetic field, even when the spin–orbit interaction is taken into account, the Hamiltonian is invariant under time reversal:

$$\mathcal{H}(\boldsymbol{r}_i, \boldsymbol{p}_i, \boldsymbol{s}_i) = \mathcal{H}(\boldsymbol{r}_i, -\boldsymbol{p}_i, -\boldsymbol{s}_i)\,. \tag{6.1.29}$$

So, besides ψ, $T\psi$ is also an eigenfunction, with the same energy. However, as we shall prove it below, when the number of electrons is odd, (6.1.27) implies the orthogonality of the two eigenfunctions – and so the state is at least doubly degenerate, as asserted by Kramers' theorem.

To prove their orthogonality, consider the scalar product of two functions f and g. The properties of the product imply

$$(K_0 f, K_0 g) = (f, g)^* = (g, f) \tag{6.1.30}$$

for the complex conjugate functions. By taking time reversal instead of complex conjugation, the involved spin factors do not spoil the relation, therefore

$$(Tf, Tg) = (g, f)\,, \tag{6.1.31}$$

in agreement with the antiunitarity of time reversal. Applying this relation to the states ψ and $T\psi$, and using (6.1.27)), we have

$$(\psi, T\psi) = (T^2\psi, T\psi) = (-1)^n(\psi, T\psi)\,. \tag{6.1.32}$$

The last equation shows that for n odd, the two states are orthogonal indeed, and so besides ψ, $T\psi$ is another independent eigenfunction, with the same energy.

6.1.5 Selection Rules

In solid-state physics the following questions are often addressed: How does the initial state of a system change under some external perturbation? What new states may arise? Quantum mechanics tells us that when the system is initially in an eigenstate $\psi_{\mathrm{i}}(\boldsymbol{r})$ of the unperturbed Hamiltonian and perturbation is given by the interaction term $\mathcal{H}_{\mathrm{int}}$ then the probability for a transition into a state with wavefunction $\psi_{\mathrm{f}}(\boldsymbol{r})$ is proportional to the absolute square of the matrix element

$$M_{\mathrm{i}\to\mathrm{f}} = \int \psi_{\mathrm{f}}^{*}(\boldsymbol{r})\mathcal{H}_{\mathrm{int}}\psi_{\mathrm{i}}(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r}\,. \tag{6.1.33}$$

For the system of electrons and for the atoms (ions), as well as the crystalline state built up of them, specification of the set of transitions that are allowed or forbidden for symmetry reasons is of particular interest. Rules that determine through symmetry considerations the set of states into which transition from a particular state is possible under a given perturbation are called *selection rules*. (A more rigorous mathematical formulation is given in Appendix D.) Here we just note that selection rules arise because of the possibility to classify the eigenstates of the unperturbed system that are involved in the transition according to the irreducible representations of the symmetry group of the unperturbed Hamiltonian. Since transition probabilities have to be invariant under symmetry operations, transitions are allowed only between states with specific symmetry properties. As an application, below we shall examine the restrictions imposed by translational symmetry.

6.2 Consequences of Translational Symmetry

When studying crystalline systems one is frequently concerned with the determination of the energy eigenvalues and eigenstates of a lattice-periodic Hamiltonian $\mathcal{H}$: just like the crystal itself, the Hamiltonian must also be invariant under translations through Bravais lattice vectors in an ideal crystal. In other words

$$\mathcal{H}(\boldsymbol{r}+\boldsymbol{t}_n) = \mathcal{H}(\boldsymbol{r}) \tag{6.2.1}$$

is required, where $\boldsymbol{t}_n$ is a translation vector of the form (5.1.1). The explicit form of the Hamiltonian is irrelevant now – therefore the following results will equally apply to electron states, phonon states of lattice vibrations, or magnetic excitations.

Rigorously speaking, translation symmetry is violated in finite lattices, even though invariance under discrete translations highly facilitates the mathematical description of the properties of crystalline materials. As surface phenomena can be usually ignored in the calculation of macroscopic properties of bulk materials, infinitely large samples seem to be the natural choice in theoretical studies. However, the mathematical discussion is far more complicated when the number of symmetry elements is infinite rather than finite.

A suitable choice of the boundary condition allows for a reduction to a finite number of symmetry elements without losing the advantages of strict translational invariance.

6.2.1 The Born–von Kármán Boundary Condition

Max Born[5] and T. von Kármán (1912) proposed a particular boundary condition for the description of finite samples that opens the way to exploiting the consequences of translational symmetry. The *periodic* or *Born–von Kármán boundary condition* can be formulated in two ways:

1. Consider a crystal that is built up of identical groups of atoms repeated N_1 times along direction $\boldsymbol{a}_1$, N_2 times along direction $\boldsymbol{a}_2$, and N_3 times along direction $\boldsymbol{a}_3$. Imposing periodic (Born–von Kármán) boundary condition amounts to formally identifying points on opposite faces of the crystal with each other. Thus translation through $N_j\boldsymbol{a}_j$ $(j = 1, 2, 3)$ takes any point $\boldsymbol{r}$ of the lattice into itself. Intuitively this can be imagined as folding one end of the sample back to the other. One-dimensional chains with two free ends are thus transformed into rings, and two-dimensional finite sheets into tori. No intuitive picture can be given for three-dimensional crystals – therefore preference is given to the second formulation in this case.
2. The periodic boundary condition can be regarded as if the originally finite, parallelepiped-shaped sample were translated through all integral linear combinations of the vectors $N_j\boldsymbol{a}_j$. The obtained pattern fills the entire space, and the point at $\boldsymbol{r}$ is rigorously equivalent to all points

$$\boldsymbol{r}' = \boldsymbol{r} + \sum_{j=1}^{3} p_j N_j \boldsymbol{a}_j \,, \tag{6.2.2}$$

where the p_j are integers.

The two formulations of the Born–von Kármán boundary condition are illustrated in Fig. 6.3 for a one-dimensional chain of atoms.

When only nearest neighbors interact, the original chain of N atoms contains $N - 1$ bonds – one less than the ring. The energy spectra of the two systems are thus not identical. Deviations are expected to be on the order of $1/N$, and are thus negligible for macroscopic samples. When the chain with free ends is replaced by a ring, other choices for the additional bond are also possible. For example, when studying a system of spins localized on atoms, in some cases it might be advantageous to choose the interaction between the two spins at the two ends to differ in sign or a phase factor from the

[5] Max Born (1882–1970) was awarded the Nobel Prize in 1954 "for his fundamental research in quantum mechanics, especially for his statistical interpretation of the wavefunction".

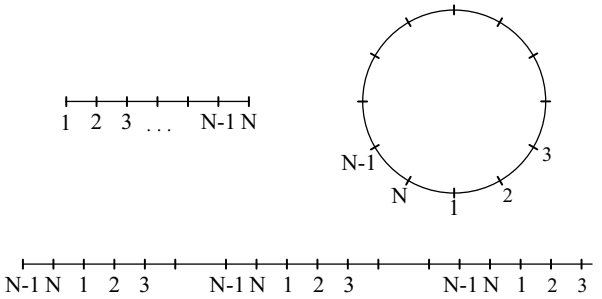


Fig. 6.3. Two different formulations of the Born–von Kármán boundary condition on a finite one-dimensional chain of atoms. The chain is either wrapped into a ring or repeated indefinitely. Consequences of translational symmetry can be exploited in both cases

interaction within the chain. This boundary condition is called *antiperiodic* or *twisted*. For electrons moving along a one-dimensional chain, the choice of periodic (antiperiodic) boundary condition is more practical when the number of fermions is odd (even).[6]

The restriction imposed by the Born–von Kármán boundary conditions can be formulated in terms of the quantum mechanical wavefunction as

$$\psi(\boldsymbol{r}) = \psi(\boldsymbol{r} + N_j \boldsymbol{a}_j)\,, \qquad j = 1, 2, 3\,. \tag{6.2.3}$$

Its counterpart for antiperiodic boundary conditions is

$$\psi(\boldsymbol{r}) = -\psi(\boldsymbol{r} + N_j \boldsymbol{a}_j)\,. \tag{6.2.4}$$

6.2.2 Bloch's Theorem

The lattice-periodic character of the Hamiltonian may not be exhibited by the wavefunction, which is not observable in itself. However, *Bloch's theorem*[7] applies to the solutions of the Schrödinger equation with a lattice-periodic potential. A general formulation of the theorem reads: *Solutions to the Schrödinger equation for a lattice-periodic Hamiltonian can be characterized by a vector quantum number* $\boldsymbol{k}$ *defined in the primitive cell of the reciprocal lattice. This "wave vector" governs the behavior under translations in the sense that values of* $\psi_{\boldsymbol{k}}(\boldsymbol{r})$ *– i.e., the wavefunction associated with the quantum number* $\boldsymbol{k}$ *– taken in different primitive cells differ by a simple phase factor:*

$$\boxed{\psi_{\boldsymbol{k}}(\boldsymbol{r} + \boldsymbol{t}_n) = \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{t}_n}\psi_{\boldsymbol{k}}(\boldsymbol{r})\,.} \tag{6.2.5}$$

[6] This must be so, since if the number of electrons is even, moving an electron from one end of the chain to the other equivalent end requires exchanging it with an odd number of electrons.

[7] F. Bloch, 1928. The one-dimensional version of this theorem is also known as *Floquet's theorem* in other branches of physics and mathematics.

Alternatively, as it will be discussed in Chapter 17, Bloch's theorem states that the eigenfunctions of a lattice-periodic Hamiltonian may be written in the form

$$\psi_{\boldsymbol{k}}(\boldsymbol{r}) = \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} u_{\boldsymbol{k}}(\boldsymbol{r})\,, \tag{6.2.6}$$

where $u_{\boldsymbol{k}}(\boldsymbol{r})$ is lattice periodic. As we shall see, $\boldsymbol{k}$ can take only well-defined discrete values determined by the specific boundary condition.

To prove Bloch's theorem, we shall introduce the linear operator $O(\boldsymbol{t}_n)$ of the translation through $\boldsymbol{t}_n$. By definition, for any function $f(\boldsymbol{r})$

$$O(\boldsymbol{t}_n) f(\boldsymbol{r}) = f(\boldsymbol{r} - \boldsymbol{t}_n)\,. \tag{6.2.7}$$

First, we shall demonstrate that $O(\boldsymbol{t}_n)$ commutes with the lattice periodic Hamiltonian $\mathcal{H}(\boldsymbol{r})$. By operating $O(\boldsymbol{t}_n)$ on $\mathcal{H}(\boldsymbol{r})f(\boldsymbol{r})$ and making use of (6.2.1) and (6.2.7),

$$\begin{aligned} O(\boldsymbol{t}_n)\mathcal{H}(\boldsymbol{r})f(\boldsymbol{r}) &= \mathcal{H}(\boldsymbol{r} - \boldsymbol{t}_n) f(\boldsymbol{r} - \boldsymbol{t}_n) \\ &= \mathcal{H}(\boldsymbol{r}) f(\boldsymbol{r} - \boldsymbol{t}_n) = \mathcal{H}(\boldsymbol{r}) O(\boldsymbol{t}_n) f(\boldsymbol{r}) \end{aligned} \tag{6.2.8}$$

is obtained, that is, the Hamiltonian commutes with any operator associated with a translation through a lattice vector:

$$O(\boldsymbol{t}_n)\mathcal{H}(\boldsymbol{r}) = \mathcal{H}(\boldsymbol{r}) O(\boldsymbol{t}_n)\,. \tag{6.2.9}$$

As $\mathcal{H}(\boldsymbol{r})$ and $O(\boldsymbol{t}_n)$ commute, they can be diagonalized simultaneously. The obtained wavefunctions are then denoted by $\psi_\alpha(\boldsymbol{r})$, and the eigenvalues by ε_α and $t_\alpha(\boldsymbol{t}_n)$,

$$\begin{aligned} \mathcal{H}(\boldsymbol{r})\psi_\alpha(\boldsymbol{r}) &= \varepsilon_\alpha \psi_\alpha(\boldsymbol{r})\,, \\ O(\boldsymbol{t}_n)\psi_\alpha(\boldsymbol{r}) &= t_\alpha(\boldsymbol{t}_n)\psi_\alpha(\boldsymbol{r})\,. \end{aligned} \tag{6.2.10}$$

Here the label α serves only to distinguish the states. Its possible values will be determined later.

The succession of two translations through $\boldsymbol{t}_n$ and $\boldsymbol{t}_{n'}$ is the same as a single translation through $\boldsymbol{t}_n + \boldsymbol{t}_{n'}$:

$$O(\boldsymbol{t}_n)\, O(\boldsymbol{t}_{n'}) = O(\boldsymbol{t}_n + \boldsymbol{t}_{n'})\,. \tag{6.2.11}$$

This equation defines group multiplication in the group of translations. As the result is independent of the order of the two translations, group multiplication is commutative,

$$O(\boldsymbol{t}_n)\, O(\boldsymbol{t}_{n'}) = O(\boldsymbol{t}_{n'})\, O(\boldsymbol{t}_n)\,, \tag{6.2.12}$$

and so translations form a commutative or Abelian group. Irreducible representations of Abelian groups are known to be one-dimensional, thus the quantum number α also serves to index them.

When the translation operator acts on the basis function of the one-dimensional representation of index α, (6.2.11) implies

$$t_\alpha(\boldsymbol{t}_n)t_\alpha(\boldsymbol{t}_{n'}) = t_\alpha(\boldsymbol{t}_n + \boldsymbol{t}_{n'}) \tag{6.2.13}$$

for the eigenvalues. Translation through $\boldsymbol{t}_n$ can be built up of translations through primitive vectors, in the form given by (5.1.1). Thus the translation operator $O(\boldsymbol{t}_n)$ can also be expressed in terms of the operators that correspond to the elementary translation vectors,

$$O(\boldsymbol{t}_n) = [O(\boldsymbol{a}_1)]^{n_1}\,[O(\boldsymbol{a}_2)]^{n_2}\,[O(\boldsymbol{a}_3)]^{n_3}\,. \tag{6.2.14}$$

Consequently, the eigenvalues of the translation operator satisfy the equation

$$t_\alpha(\boldsymbol{t}_n) = [t_\alpha(\boldsymbol{a}_1)]^{n_1}\,[t_\alpha(\boldsymbol{a}_2)]^{n_2}\,[t_\alpha(\boldsymbol{a}_3)]^{n_3}\,. \tag{6.2.15}$$

The eigenvalues $t_\alpha(\boldsymbol{t}_n)$ can be determined from the requirement that the boundary conditions imposed on the wavefunctions should also be satisfied. For a general parallelepiped-shaped sample the Born–von Kármán boundary conditions (6.2.3) may be written as

$$\psi(\boldsymbol{r} - N_j\boldsymbol{a}_j) = \psi(\boldsymbol{r})\,, \qquad j = 1,2,3\,. \tag{6.2.16}$$

This implies

$$O(N_1\boldsymbol{a}_1) = O(N_2\boldsymbol{a}_2) = O(N_3\boldsymbol{a}_3) = 1\,, \tag{6.2.17}$$

leading to

$$[t_\alpha(\boldsymbol{a}_1)]^{N_1} = [t_\alpha(\boldsymbol{a}_2)]^{N_2} = [t_\alpha(\boldsymbol{a}_3)]^{N_3} = 1\,. \tag{6.2.18}$$

The solutions of these equations are the roots of unity,

$$t_\alpha(\boldsymbol{a}_j) = \mathrm{e}^{-2\pi\mathrm{i}p_j/N_j}\,, \tag{6.2.19}$$

where p_j is an integer. The N_j different roots are most simply obtained by restricting p_j to the range

$$0 \le p_j \le N_j - 1\,. \tag{6.2.20}$$

Substituting the above form of $\boldsymbol{t}_\alpha(\boldsymbol{a}_i)$ into (6.2.15) gives

$$t_\alpha(\boldsymbol{t}_n) = \exp[-2\pi\mathrm{i}(n_1p_1/N_1 + n_2p_2/N_2 + n_3p_3/N_3)]\,. \tag{6.2.21}$$

This eigenvalue can thus be expressed in terms of the three quantum numbers p_1, p_2, p_3. The precise specification of the state may require further quantum numbers that appear in the wavefunction and the energy eigenvalues – but not in the eigenvalues of the translation operator. Below, we shall focus on the triplet p_1, p_2, p_3 and ignore these further quantum numbers; their importance will transpire only later.

The expression for $t_\alpha(\boldsymbol{t}_n)$ may be written in a more concise form using the notation

$$\boldsymbol{k} = \frac{p_1}{N_1}\boldsymbol{b}_1 + \frac{p_2}{N_2}\boldsymbol{b}_2 + \frac{p_3}{N_3}\boldsymbol{b}_3\,, \tag{6.2.22}$$

where $\boldsymbol{b}_1$, $\boldsymbol{b}_2$, and $\boldsymbol{b}_3$ are the primitive vectors of the reciprocal lattice. Because of the restrictions on p_j, the $\boldsymbol{k}$ are within the primitive cell spanned by the vectors $\boldsymbol{b}_i$. When the p_j run over their entire range, the resulting set of vectors fill this primitive cell evenly.

Using the multiplication rule (5.2.12) for the product of a direct- and a reciprocal-lattice vector, the expression in (6.2.21) can be written as

$$t_{\boldsymbol{k}}(\boldsymbol{t}_n) = \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{t}_n} \,. \tag{6.2.23}$$

Returning to (6.2.10), the solutions $\psi_{\boldsymbol{k}}(\boldsymbol{r})$ to the Schrödinger equation are found to have the following property:

$$\psi_{\boldsymbol{k}}(\boldsymbol{r}+\boldsymbol{t}_n) = O(-\boldsymbol{t}_n)\psi_{\boldsymbol{k}}(\boldsymbol{r}) = t_{\boldsymbol{k}}(-\boldsymbol{t}_n)\psi_{\boldsymbol{k}}(\boldsymbol{r}) = \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{t}_n}\psi_{\boldsymbol{k}}(\boldsymbol{r}) \,. \tag{6.2.24}$$

This completes the proof of Bloch's theorem. As a by-product, we have also derived the allowed values of the quantum number specifying the behavior under translations, the wave vector (also called wave-number vector) $\boldsymbol{k}$.

The phase factor contains the scalar product of the discrete translation vector and the wave vector. For continuous translations a Taylor expansion and subsequent summation of the wavefunction at the position shifted through $\boldsymbol{a}$, $\psi(\boldsymbol{r}+\boldsymbol{a})$ would lead to

$$\begin{aligned}\psi(\boldsymbol{r}+\boldsymbol{a}) &= \psi(\boldsymbol{r}) + \sum_{\alpha} \frac{\partial}{\partial r_\alpha}\psi(\boldsymbol{r})a_\alpha + \tfrac{1}{2}\sum_{\alpha\beta}\frac{\partial^2}{\partial r_\alpha \partial r_\beta}\psi(\boldsymbol{r})a_\alpha a_\beta + \ldots \\ &= \exp\Big(\frac{\partial}{\partial \boldsymbol{r}}\cdot\boldsymbol{a}\Big)\psi(\boldsymbol{r}) \,. \end{aligned} \tag{6.2.25}$$

This can be expressed in terms of the canonical momentum $\boldsymbol{p}$ as

$$\psi(\boldsymbol{r}+\boldsymbol{a}) = \mathrm{e}^{\mathrm{i}\boldsymbol{p}\cdot\boldsymbol{a}/\hbar}\psi(\boldsymbol{r}) \,. \tag{6.2.26}$$

This formula suggests that $\hbar\boldsymbol{k}$ plays the same role in discrete translations as momentum in continuous translations. This analogy justifies calling $\hbar\boldsymbol{k}$ the *crystal momentum*. Nevertheless it should be stressed that the state $\psi_{\boldsymbol{k}}(\boldsymbol{r})$ is not necessarily an eigenstate of the momentum operator, and its momentum is not necessarily $\hbar\boldsymbol{k}$.

6.2.3 Equivalent Wave Vectors

The number of different $\boldsymbol{k}$ vectors allowed by the boundary condition is $N = N_1 \times N_2 \times N_3$, i.e., the number of primitive cells in the finite crystal. The choice of the wave vectors is, however, not unique. In the previous subsection the quantum numbers p_i (that appeared in the definition of $\boldsymbol{k}$) were chosen according to the condition (6.2.20). This particular choice implied that the wave numbers fill a primitive cell of the reciprocal lattice.[8] However, this

[8] Throughout the present subsection "primitive cell" should be understood as "primitive cell spanned by the primitive vectors of the reciprocal lattice".

choice is not unique. As the $\boldsymbol{k}$s were determined from the requirement that (6.2.18) should be satisfied, the same eigenvalues are found when p_j is replaced by $p_j + h_j N_j$ in (6.2.19) – where h_j is an arbitrary integer. The behavior of the system under translations is then characterized by the vector

$$\begin{aligned}\boldsymbol{k}' &= \frac{p_1 + h_1 N_1}{N_1}\boldsymbol{b}_1 + \frac{p_2 + h_2 N_2}{N_2}\boldsymbol{b}_2 + \frac{p_3 + h_3 N_3}{N_3}\boldsymbol{b}_3 \\ &= \boldsymbol{k} + h_1\boldsymbol{b}_1 + h_2\boldsymbol{b}_2 + h_3\boldsymbol{b}_3\end{aligned} \tag{6.2.27}$$

instead of $\boldsymbol{k}$. The two differ by a reciprocal-lattice vector,

$$\boldsymbol{k}' = \boldsymbol{k} + \boldsymbol{G}\,. \tag{6.2.28}$$

It is readily seen that it makes no difference in Bloch's theorem whether the phase factor is determined using $\boldsymbol{k}$ or $\boldsymbol{k}'$. For any translation through a lattice vector,

$$t_{\boldsymbol{k}}(\boldsymbol{t}_n) = t_{\boldsymbol{k}'}(\boldsymbol{t}_n)\,. \tag{6.2.29}$$

Thus $\boldsymbol{k}$ and $\boldsymbol{k}'$ – which differ by a vector of the reciprocal lattice – can be equivalently used as a quantum number of the states. In this respect, the two vectors are equivalent.

As it was mentioned in Chapter 5, the object obtained with Dirichlet's construction in the reciprocal lattice is called the *Brillouin zone*. By appropriately cutting the primitive cell and translating the parts through suitably chosen reciprocal-lattice vectors, the Brillouin zone is totally covered without overlap in an unambiguous way. The parallelepiped-shaped primitive cell, the Brillouin zone, and the parts that have to be translated to make the two overlap are shown in Fig. 6.4 for a two-dimensional reciprocal lattice.

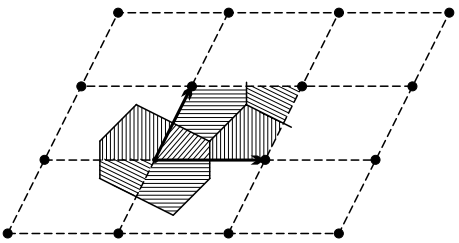


Fig. 6.4. Equivalence of the primitive cell spanned by the primitive vectors and the Brillouin zone of a two-dimensional reciprocal lattice. Regions with identical shading are displaced through some lattice vector

Since each vector $\boldsymbol{k}$ within the primitive cell can be unambiguously associated with a wave vector inside the Brillouin zone, the states of the lattice-periodic system can be specified using vectors $\boldsymbol{k}$ in the Brillouin zone. More often than not, this path is taken in solid-state physics. The reason for this

is related to symmetries. As will be demonstrated later, the energy spectrum of states characterized by the wave vector $\boldsymbol{k}$ displays the same symmetries as the crystal itself. On the other hand, as it has been mentioned, the primitive cell spanned by the primitive vectors does not usually exhibit the symmetries of the crystal, whereas the Brillouin zone does (similarly to the Wigner–Seitz cell of the direct lattice). Thus the symmetries of the energy spectrum are more transparent when the latter is represented in the Brillouin zone.

6.2.4 Conservation of Crystal Momentum

Noether's theorem[9] establishes that invariance under symmetry transformations usually implies conservation laws. Spatial translations are related to momentum conservation; time translations to energy conservation; rotational symmetry to angular-momentum conservation; and gauge symmetry to charge conservation.

Crystalline systems are not invariant under arbitrary translations, therefore, strictly speaking, momentum is not conserved. However, as mentioned above, because of discrete translational symmetry, crystalline states may be characterized by a wave vector $\boldsymbol{k}$, and the quantity $\hbar\boldsymbol{k}$ plays a role similar to momentum. We shall prove that although no strict conservation law is valid for $\boldsymbol{k}$, discrete symmetries nevertheless imply a conservation law of rather limited validity. For this reason the crystal momentum $\hbar\boldsymbol{k}$ is often called the *quasimomentum*; the same name is often applied to $\boldsymbol{k}$ itself in the literature.

In quantum mechanics, the time derivative of the translation operator is given by

$$\frac{\mathrm{d}O(\boldsymbol{t}_n)}{\mathrm{d}t} = \frac{\mathrm{i}}{\hbar}\left[\mathcal{H}, O(\boldsymbol{t}_n)\right]. \tag{6.2.30}$$

Since the Hamiltonian commutes with the translation operator in crystalline systems, we have

$$\frac{\mathrm{d}O(\boldsymbol{t}_n)}{\mathrm{d}t} = 0\,. \tag{6.2.31}$$

Consequently, the wave vector $\boldsymbol{k}$ that characterizes the transformation of the wavefunction under translations is independent of time.

Suppose that the wavefunction of the system can be written as the product of the wavefunctions of r subsystems,

$$\psi_{\boldsymbol{k}} = \phi_{\boldsymbol{k}_1}\phi_{\boldsymbol{k}_2}\ldots\phi_{\boldsymbol{k}_r}\,, \tag{6.2.32}$$

where the behavior of the subsystems under translations is characterized by a set of wave vectors $\boldsymbol{k}_l$ (quasimomenta $\hbar\boldsymbol{k}_l$), that is,

$$O(\boldsymbol{t}_n)\phi_{\boldsymbol{k}_l}(\boldsymbol{r}) = \mathrm{e}^{-\mathrm{i}\boldsymbol{k}_l\cdot\boldsymbol{t}_n}\phi_{\boldsymbol{k}_l}(\boldsymbol{r})\,. \tag{6.2.33}$$

Then the behavior of the whole system under translations is governed by

[9] E. Noether, 1918.

$$O(\boldsymbol{t}_n)\psi_{\boldsymbol{k}}(\boldsymbol{r}) = \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{t}_n}\psi_{\boldsymbol{k}}(\boldsymbol{r})\,, \tag{6.2.34}$$

where the wave vector $\boldsymbol{k}$ of the whole system is the sum of the wave vectors of its subsystems:

$$\boldsymbol{k} = \sum_{l=1}^{r} \boldsymbol{k}_l\,. \tag{6.2.35}$$

This is true even if the complete wavefunction of the system has to be constructed as the symmetrized or antisymmetrized product of the wavefunctions of the subsystems.

Microscopic processes that occur in the crystal – e.g., collisions between electrons – may change the wavefunctions,[10] and thus the quasimomenta of the subsystems. Assume that, as a result of such an internal collision, the initial state

$$\psi_{\boldsymbol{k}} = \phi_{\boldsymbol{k}_1}\phi_{\boldsymbol{k}_2}\ldots\phi_{\boldsymbol{k}_r} \tag{6.2.36}$$

of the system changes into the final state

$$\psi_{\boldsymbol{k}'} = \phi_{\boldsymbol{k}'_1}\phi_{\boldsymbol{k}'_2}\ldots\phi_{\boldsymbol{k}'_{r'}}\,, \tag{6.2.37}$$

where

$$\boldsymbol{k} = \sum_{l=1}^{r} \boldsymbol{k}_l\,, \qquad \boldsymbol{k}' = \sum_{l=1}^{r'} \boldsymbol{k}'_l\,. \tag{6.2.38}$$

If the system is not subject to any external perturbation, then constancy of $\boldsymbol{k}$ in time means

$$\boldsymbol{k} \equiv \boldsymbol{k}'\,. \tag{6.2.39}$$

Instead of equality – which would rigorously hold for continuous translational symmetry – only the equivalence ($\equiv$) of the initial and final wave vectors is guaranteed by discrete translational symmetry, that is, the two may differ by a reciprocal-lattice vector:

$$\boldsymbol{k}' = \boldsymbol{k} + \boldsymbol{G}\,. \tag{6.2.40}$$

This is because only the phase factors that appear in the wavefunctions have to be equal – that is, for a translation through an arbitrary lattice vector $\boldsymbol{t}_n$ only the equality

$$\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{t}_n} = \mathrm{e}^{-\mathrm{i}\boldsymbol{k}'\cdot\boldsymbol{t}_n} \tag{6.2.41}$$

is required. This is also met by equivalent wave vectors.

In vacuum, electrons interact via Coulomb repulsion. The situation is more complicated in solids. As we shall see in Volume 3, the effects of other electrons can be taken into account using some kind of modified, effective interaction between electrons. However, just like Coulomb interaction, this effective interaction is a pair interaction: it leads to elementary processes in which the state of only two electrons change at the same time. To put it otherwise: in an elementary process two electrons with quasimomenta $\hbar\boldsymbol{k}_1$ and $\hbar\boldsymbol{k}_2$ are scattered into a state with quasimomenta $\hbar\boldsymbol{k}'_1$ and $\hbar\boldsymbol{k}'_2$. This scattering process is illustrated in Fig. 6.5.

[10] Note that even in thermal equilibrium the system is not in a pure state.

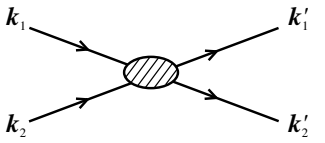


Fig. 6.5. Diagrammatic representation of the interaction between two electrons, with the initial and final wave vectors of the particles

Strict momentum conservation applies to all scattering processes among free electrons. According to the foregoing, this requirement is replaced by

$$\boldsymbol{k}_1' + \boldsymbol{k}_2' = \boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{G} \tag{6.2.42}$$

for electrons moving in a crystal lattice.

The interaction of electrons with lattice vibrations may be interpreted in terms of the creation (emission) and annihilation (absorption) of a quantum of lattice vibration – a phonon. Such processes are shown in Fig. 6.6, where solid lines represent electrons and wavy lines the quanta of lattice vibrations emitted or absorbed in the interaction.

Fig. 6.6. Diagrammatic representation of the emission and absorption of quanta of lattice vibrations by electrons. Solid lines are electrons, wavy lines are lattice vibrations

In a process where a single electron absorbs a phonon of wave vector $\boldsymbol{q}$, the electron is scattered into a state with wave vector

$$\boldsymbol{k}' = \boldsymbol{k} + \boldsymbol{q} + \boldsymbol{G}\,. \tag{6.2.43}$$

In processes where phonons are emitted

$$\boldsymbol{k}' + \boldsymbol{q} = \boldsymbol{k} + \boldsymbol{G}\,. \tag{6.2.44}$$

Processes in which $\boldsymbol{G} = 0$ are called *normal processes* (N-processes), while those in which $\boldsymbol{G} \neq 0$ are called *umklapp processes* (U-processes).[11] This distinction is important because normal and umklapp processes may give essentially different contributions to the macroscopic properties of the crystal. For

[11] R. E. Peierls, 1929. The English equivalent of *umklapp process* is *flip-over process*, however this is hardly used in the literature.

example, thermal conductivity cannot be understood without taking umklapp processes into account. On the other hand, there are some physical properties of solids that may be interpreted without recourse to the discreteness of the lattice. When the lattice model of a crystal is replaced by a continuous one, the size of the primitive cell goes to zero, and the vectors of the reciprocal lattice become infinitely large. Umklapp processes are then negligible, and strict momentum conservation is recovered.

The conservation of quasimomentum may be applied to the scattering of an incident beam of photons, neutrons, or electrons by crystalline samples. Assume that the incident beam is described by a plane wave, i.e., the incoming particles have a well-defined momentum (or wave number). Due to the discrete translational symmetry of the sample, the initial wave vector $\boldsymbol{k}$ of the incoming particle may change by a reciprocal-lattice vector in elastic scattering processes – i.e, when the state of the sample does not change. Thus the outgoing beam is made up of particles with wave vector

$$\boldsymbol{k}' = \boldsymbol{k} + \boldsymbol{G}\,, \qquad |\boldsymbol{k}'| = |\boldsymbol{k}| \tag{6.2.45}$$

or momentum $\hbar\boldsymbol{k}'$. As we shall see in Chapter 8, this constraint is just the *Laue condition* of diffraction.

In inelastic processes a quantum of lattice vibration of wave vector $\boldsymbol{q}$ may be created or absorbed in the interaction between the incoming particles and the solid. Then condition

$$\boldsymbol{k}' = \boldsymbol{k} + \boldsymbol{q} + \boldsymbol{G} \tag{6.2.46}$$

needs to be satisfied in processes with phonon emission and

$$\boldsymbol{k}' = \boldsymbol{k} - \boldsymbol{q} + \boldsymbol{G} \tag{6.2.47}$$

in processes with phonon absorption. The conservation of quasimomentum can be written similarly for other processes, e.g., scattering or absorption of light (photons) by a system of electrons.

The foregoing discussion applies to perfect crystals that are free of impurities and defects. Such imperfections are always present in real crystals, nevertheless as long as their concentration is low, the probability that scattering occurs on them is negligible, and so the above selection rules can be used. Because of the periodic boundary condition, the above considerations are valid for states within the sample, far from its surface. In processes where scattering occurs on surface states, only the component of the wave vector parallel to the surface ($\boldsymbol{k}_{\parallel}$) is conserved, the perpendicular component ($\boldsymbol{k}_{\perp}$) is not.

6.2.5 Symmetry Properties of Energy Eigenstates

As we have seen above, microscopic states within a crystalline material can be characterized by the wave vector $\boldsymbol{k}$. However, nothing has been said of the

energy $\varepsilon_{\boldsymbol{k}}$ of the states yet. We shall show that whatever the explicit form of the Hamiltonian, the symmetry properties of the system appear in $\varepsilon_{\boldsymbol{k}}$, too.

Suppose that a solution to the Schrödinger equation

$$\mathcal{H}\psi_{\boldsymbol{k}}(\boldsymbol{r}) = \varepsilon_{\boldsymbol{k}}\psi_{\boldsymbol{k}}(\boldsymbol{r}) \tag{6.2.48}$$

is known. This may be an electronic state, a vibrational state of the lattice, or a magnetic excitation. Let us now examine what happens when this state is acted upon by a symmetry element of the space group of the crystal, $\{\alpha|\boldsymbol{t}_n\}$, which is composed of a rotation α and a translation $\boldsymbol{t}_n$.

If the system is not subject to any external perturbation, then the symmetry elements of the space group of the crystal leave the Hamiltonian invariant, and it commutes with these symmetry operations, that is,

$$O(\{\alpha|\boldsymbol{t}_n\})\mathcal{H} = \mathcal{H}O(\{\alpha|\boldsymbol{t}_n\})\,. \tag{6.2.49}$$

Acting on the Schrödinger equation (6.2.48) with the symmetry operator and using the previous commutation relation we have

$$\mathcal{H}O(\{\alpha|\boldsymbol{t}_n\})\psi_{\boldsymbol{k}}(\boldsymbol{r}) = \varepsilon_{\boldsymbol{k}}O(\{\alpha|\boldsymbol{t}_n\})\psi_{\boldsymbol{k}}(\boldsymbol{r})\,. \tag{6.2.50}$$

Apparently, in addition to $\psi_{\boldsymbol{k}}(\boldsymbol{r})$, $O(\{\alpha|\boldsymbol{t}_n\})\psi_{\boldsymbol{k}}(\boldsymbol{r})$ is also an eigenfunction of the Schrödinger equation, with the same energy eigenvalue $\varepsilon_{\boldsymbol{k}}$. Had we made recourse to Wigner's theorem, we could have written down this result immediately.

Now let us examine the behavior of the wavefunction $O(\{\alpha|\boldsymbol{t}_n\})\psi_{\boldsymbol{k}}(\boldsymbol{r})$ under translations – in particular, the action of the translation operator $O(\boldsymbol{t}_m) = O(\{E|\boldsymbol{t}_m\})$ on this state. Using the relation

$$O(\{E|\boldsymbol{t}_m\})O(\{\alpha|\boldsymbol{t}_n\}) = O(\{\alpha|\boldsymbol{t}_n\})O(\{E|\alpha^{-1}\boldsymbol{t}_m\})\,, \tag{6.2.51}$$

which is a straightforward consequence of the multiplication rule for the elements of the space group, we have

$$\begin{aligned} O(\boldsymbol{t}_m)O(\{\alpha|\boldsymbol{t}_n\})\psi_{\boldsymbol{k}}(\boldsymbol{r}) &= O(\{\alpha|\boldsymbol{t}_n\})O(\alpha^{-1}\boldsymbol{t}_m)\psi_{\boldsymbol{k}}(\boldsymbol{r}) \\ &= O(\{\alpha|\boldsymbol{t}_n\})\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot(\alpha^{-1}\boldsymbol{t}_m)}\psi_{\boldsymbol{k}}(\boldsymbol{r}) \\ &= O(\{\alpha|\boldsymbol{t}_n\})\mathrm{e}^{-\mathrm{i}(\alpha\boldsymbol{k})\cdot\boldsymbol{t}_m}\psi_{\boldsymbol{k}}(\boldsymbol{r}) \\ &= \mathrm{e}^{-\mathrm{i}(\alpha\boldsymbol{k})\cdot\boldsymbol{t}_m}O(\{\alpha|\boldsymbol{t}_n\})\psi_{\boldsymbol{k}}(\boldsymbol{r})\,. \end{aligned} \tag{6.2.52}$$

The wave vector of the transformed wavefunction is thus $\alpha\boldsymbol{k}$. However, as the energy of the state remains the same, we have

$$\boxed{\varepsilon_{\alpha\boldsymbol{k}} = \varepsilon_{\boldsymbol{k}}\,.} \tag{6.2.53}$$

It can be established in complete generality that $\varepsilon_{\boldsymbol{k}}$ possesses the full symmetry of the point group.[12]

[12] Note that in this derivation we did not have to assume that $\boldsymbol{t}_n$ in $\{\alpha|\boldsymbol{t}_n\}$ is a translation vector of the lattice, thus our result is also valid in the case when the space group contains glide reflections and/or screw rotations.

In later chapters dealing with electrons, phonons, and magnons, one of our most important tasks will be the determination of the energy $\varepsilon_{\boldsymbol{k}}$ for all possible values of $\boldsymbol{k}$ – as this provides the dispersion relation of the excitation. By exploiting the symmetries of $\varepsilon_{\boldsymbol{k}}$, the energy values do not need to be determined for each point of the full Brillouin zone separately: calculations or measurements can be restricted to a small portion of the zone. Energies for any other $\boldsymbol{k}$ can then be established using symmetry considerations.

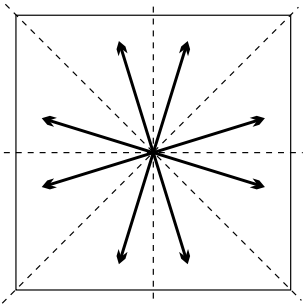


Fig. 6.7. Wave vectors with the same energy in the Brillouin zone of a two-dimensional square lattice

Figure 6.7 clearly shows that in a two-dimensional square crystal calculations have to be performed in one-eighth of the Brillouin zone, as symmetry operations relate any general point $\boldsymbol{k} = (k_x, k_y)$ to seven others. In a three-dimensional cubic crystal a general point (k_x, k_y, k_z) of the Brillouin zone is related to 47 others by the 48 symmetry operations of a cube, as it can be read off from Tables 5.1 and 5.4. Thus in this case the spectrum needs to be determined on 1/48 of the Brillouin zone.

Next, let us consider crystals with inversion centers. Inversion transforms the variable $\boldsymbol{r}$ into $-\boldsymbol{r}$, and a state with wave vector $\boldsymbol{k}$ into another with wave vector $-\boldsymbol{k}$. However, because of its axial vector character spin is left invariant, and so

$$\varepsilon_{\boldsymbol{k},\sigma} = \varepsilon_{-\boldsymbol{k},\sigma} \tag{6.2.54}$$

For spinless excitations – e.g., excitations of lattice vibrations – the relation

$$\varepsilon_{\boldsymbol{k}} = \varepsilon_{-\boldsymbol{k}} \tag{6.2.55}$$

is valid more generally, even in crystals without an inversion center. This is the consequence of invariance under time reversal, as in this particular case time reversal is equivalent to complex conjugation. By taking the complex conjugate (cc) of the Schrödinger equation, and using the Hermiticity of the Hamiltonian,

$$\mathcal{H}\psi^*_{\boldsymbol{k}}(\boldsymbol{r}) = \varepsilon_{\boldsymbol{k}}\psi^*_{\boldsymbol{k}}(\boldsymbol{r})\,. \tag{6.2.56}$$

The transformation properties of wavefunctions under translations imply that complex conjugation of the wavefunction takes a state with quasimomentum $\boldsymbol{k}$ into one with quasimomentum $-\boldsymbol{k}$. This means that for spinless excitations the excitation energy is always an even function of the wave number.

The situation is more complicated for electrons. If spin–orbit interactions are taken into account but no magnetic field is present then the Hamiltonian

$$\mathcal{H} = -\frac{\hbar^2}{2m_\mathrm{e}}\boldsymbol{\nabla}^2 + U(\boldsymbol{r}) + \frac{\hbar}{2m_\mathrm{e}^2c^2}\frac{1}{r}\frac{\mathrm{d}U(r)}{\mathrm{d}r}(\boldsymbol{r}\times\boldsymbol{p})\cdot\boldsymbol{s}\,, \tag{6.2.57}$$

which contains the term given in (3.1.33), is invariant under time reversal. Therefore the time reversed of each eigenfunction is also an eigenfunction, with the same energy. However, the presence of spin–orbit interaction no longer permits us to index the eigenstates with the spin z component: the latter is not a good quantum number any more. Time reversal changes the sign of orbital angular momentum and spin alike: $\boldsymbol{p}\to-\boldsymbol{p}$ and $\boldsymbol{s}\to-\boldsymbol{s}$. The eigenstates $\psi_{\boldsymbol{k},\alpha}(\boldsymbol{r})$ and $\psi_{\boldsymbol{k},\beta}(\boldsymbol{r})$ of the Hamiltonian can therefore be written as linear combinations of the $m_s=\pm\frac{1}{2}$ states $|\uparrow\rangle$ and $|\downarrow\rangle$:

$$\begin{aligned}\psi_{\boldsymbol{k},\alpha}(\boldsymbol{r}) &= \chi_{\boldsymbol{k},+}(\boldsymbol{r})|\uparrow\rangle + \chi_{\boldsymbol{k},-}(\boldsymbol{r})|\downarrow\rangle\,,\\ \psi_{\boldsymbol{k},\beta}(\boldsymbol{r}) &= -\chi_{\boldsymbol{k},-}(\boldsymbol{r})|\uparrow\rangle + \chi_{\boldsymbol{k},+}(\boldsymbol{r})|\downarrow\rangle\,.\end{aligned} \tag{6.2.58}$$

These states may be chosen such that $\psi_{\boldsymbol{k},\alpha}(\boldsymbol{r})$ go over into $\psi_{\boldsymbol{k},\uparrow}(\boldsymbol{r})$ and $\psi_{\boldsymbol{k},\beta}(\boldsymbol{r})$ into $\psi_{\boldsymbol{k},\downarrow}(\boldsymbol{r})$ when the strength of the spin–orbit interaction vanishes.

Instead of (6.1.23), the time reversal operator may be written as

$$T = e^{-\mathrm{i}\pi s_y}K_0 = -\mathrm{i}\sigma_y K_0\,, \tag{6.2.59}$$

which amounts to a different choice of the phase factor. When applied to the above states,

$$T\psi_{\boldsymbol{k},\alpha}(\boldsymbol{r}) = \psi_{-\boldsymbol{k},\beta}(\boldsymbol{r}) \quad\text{and}\quad T\psi_{\boldsymbol{k},\beta}(\boldsymbol{r}) = -\psi_{-\boldsymbol{k},\alpha}(\boldsymbol{r})\,, \tag{6.2.60}$$

which leads to

$$\boxed{\varepsilon_{\boldsymbol{k},\alpha} = \varepsilon_{-\boldsymbol{k},\beta}\,, \qquad \varepsilon_{\boldsymbol{k},\beta} = \varepsilon_{-\boldsymbol{k},\alpha}\,.} \tag{6.2.61}$$

Owing to the spin–orbit interaction, the two spin states associated with the same wave vector $\boldsymbol{k}$ have different energies, however for each of them there exists another state with the same energy. This is just Kramers' theorem for one-particle electron states.

If the crystal has inversion symmetry in addition to time reversal invariance then relations (6.2.54) and (6.2.61) imply

$$\varepsilon_{\boldsymbol{k},\alpha} = \varepsilon_{\boldsymbol{k},\beta} \tag{6.2.62}$$

as well, that is, in the absence of an external magnetic field electron states show spin degeneracy, despite the explicit spin-dependence of the spin–orbit interaction. This degeneracy is lifted when inversion symmetry is broken.

In certain cases states that correspond to $\boldsymbol{k}$s on the boundary of the Brillouin zone have special properties: for certain symmetries the derivative of the dispersion curve – that is, the velocity of the particle associated with the excited state – vanishes. In the customary representation this means that dispersion curves arrive perpendicularly at the boundaries of the Brillouin zone. We shall investigate the conditions for this.

Suppose that one of the boundary planes of the Brillouin zone is transformed by a reflection symmetry of the crystal into another plane whose points are equivalent with those of the original plane – that is, the two plane faces can be transformed into one another by a translation through a reciprocal-lattice vector $\boldsymbol{G}$. This also means that this boundary is itself a mirror plane of the reciprocal lattice.

Consider a point $\boldsymbol{k}_1$ close to the boundary of the Brillouin zone, as shown in sectional view in Fig. 6.8. If reflection symmetry takes $\boldsymbol{k}_1$ into $\boldsymbol{k}_2$, the corresponding energies are equal: $\varepsilon_{\boldsymbol{k}_1} = \varepsilon_{\boldsymbol{k}_2}$. Translation of $\boldsymbol{k}_2$ through a reciprocal-lattice vector $\boldsymbol{G}$ gives $\boldsymbol{k}_3$. Because of the equivalence of these points $\varepsilon_{\boldsymbol{k}_3} = \varepsilon_{\boldsymbol{k}_2}$, and so $\varepsilon_{\boldsymbol{k}_1} = \varepsilon_{\boldsymbol{k}_3}$.

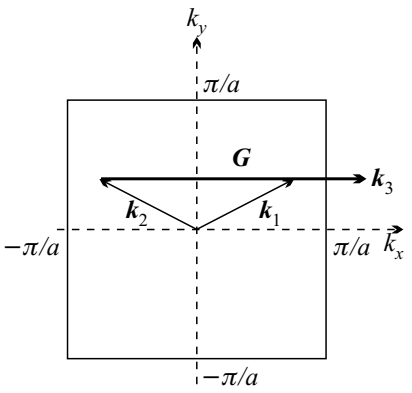


Fig. 6.8. Wave vectors that are transformed into one another by a reflection or a translation through a reciprocal-lattice vector

Unless the energy level becomes degenerate with another, the relation $\varepsilon_{\boldsymbol{k}_1} = \varepsilon_{\boldsymbol{k}_3}$ implies that as $\boldsymbol{k}$ moves along a line perpendicular to the zone boundary $\varepsilon_{\boldsymbol{k}}$ takes an extremum on the boundary. This implies that the dispersion curve is perpendicular to the boundary, i.e., its normal derivative – the velocity of the excitations – vanishes.

This requires that this piece of the zone boundary be a symmetry plane – and also that degeneracies be lifted on the boundary. If the latter requirement is not met, the previous relations imply that the derivatives of the dispersion curve are equal in magnitude but opposite in sign in the two branches. We shall see an example in the band structure of silicon and germanium.

6.3 Symmetry Breaking and Its Consequences

In the foregoing we have dealt mostly with the symmetries that appear in crystalline systems and emphasized their importance in describing physical properties. We have not investigated how this symmetry is related to the full symmetry of the system's Hamiltonian. Later, when the Hamiltonian – of the system of electrons or the lattice itself – is specified, we shall see that it possesses more symmetries than the actual physical state. The Hamiltonian is invariant under an arbitrary translation, while the crystal itself only under discrete ones. In such cases some symmetry of the Hamiltonian is said to be spontaneously broken in the physical system. Just as the presence of symmetries helps to understand the behavior of the physical system, their absence, i.e., the breaking of some symmetries of the Hamiltonian may also have important implications. Symmetry breaking may also occur when a system undergoes a transition from one phase to another. In this case symmetry considerations may be helpful in describing some features of the transition and of the new phase.

6.3.1 Symmetry Breaking in Phase Transitions

The crystalline phase, which is invariant under discrete translational symmetries, is reached through a phase transition from the liquid phase, which is invariant under arbitrary translations. Further phase transitions may occur in the crystalline state; for example, different modifications may become stable, or a magnetic or superconducting phase may appear below a certain critical temperature. In each of them, further symmetries are broken. Whether the transition itself is first- or second-order, symmetry is always broken abruptly. This is expressed by Landau's statement, which P. W. Anderson called the first theorem of condensed-matter physics: *Symmetry cannot change continuously.* A system is either invariant or not under a symmetry operation. However small the symmetry-breaking contribution in a physical quantity, if it is finite the symmetry is broken.

It should be noted that phase transition is not necessarily accompanied by a change in symmetry. A typical example for this is the transition between liquid and gas phases. However, here we are interested in the general conclusions that can be drawn from symmetry considerations provided symmetry is changed.

The basic assumption underlying the Landau theory of phase transitions[13] is that in second-order (continuous) phase transitions the phase that is stable at lower temperatures is always distinguished from the high-temperature disordered phase by a suitable order parameter. This order parameter vanishes in the disordered (usually higher-temperature) phase and appears continuously in the phase transition, which is why free energy can be expanded in its powers. We shall present a somewhat more detailed discussion in Chapter 14 in the

[13] L. D. Landau, 1937.

context of phase transitions of magnetic systems. As E. M. Lifshitz pointed out in 1941, the method provides a means to investigate which symmetries of a particular crystal might be preserved. For example, in an alloy undergoing disorder–order phase transition one may predict which sites of the crystal will be occupied by one atomic species and which sites by the other. Even more interesting is that when magnetic ordering takes place, knowledge of the symmetries of the high-temperature magnetically disordered phase permits one to predict the symmetries of the magnetically ordered states that might appear – provided the phase transition is of second order.

6.3.2 Goldstone's Theorem

The above considerations are concerned with the macroscopic properties that a state arising from symmetry breaking might have. Conversely, *Goldstone's theorem*[14] is concerned with the system's microscopic behavior. Here we give the statement without proof: *If the Hamiltonian of a system possesses some continuous symmetry that is broken in the actual physical state then – provided the interaction is not long-ranged – the energy of the lowest-lying excited states of the system are on the order of* $1/N$ *above that of the ground-state. In the* $N \to \infty$ *limit the energies of the excited states form a continuum, which is not separated by any gap from the ground state. Moreover, this continuum is similar to the energy spectrum of a system of free particles obeying Bose–Einstein statistics.*

The low-lying excited states arise from the fluctuations of the symmetry-breaking order parameter. According to the theorem the spectrum of these fluctuations can be interpreted in terms of bosonic effective particles, elementary excitations. These soft modes are called *Goldstone bosons*. The energy of the excitations is proportional to their wave number, except when the order parameter is a conserved quantity; in this case the energy is proportional to the square of the wave number.

We shall see an example for this in lattices held together by short-range forces, where breaking the continuous translational symmetry gives rise to low-energy long-wavelength acoustic phonon modes. Similarly, breaking the continuous rotational symmetry in isotropic ferro- or antiferromagnetic materials leads to low-energy spin waves. In both cases the spectrum of collective excitations is gapless. Note that the theorem states nothing about breaking a discrete symmetry of the Hamiltonian. In this case the ground state is usually separated from the continuum of excited states by a finite gap.

It is essential that Goldstone bosons do not appear for long-range forces: the spectrum of collective bosonic excitations has a gap then. When examining the excitation spectrum of an interacting electron gas we shall see that plasmons have a finite energy as Coulomb repulsion has a long range. In Wigner crystals (to be discussed in Chapter 30) two soft transverse modes appear,

[14] J. Goldstone, 1961.

however the longitudinal mode does not become soft. Goldstone bosons are absent in superconductors, too. However, in liquid helium-4, where interactions are short-ranged, the excitation spectrum is gapless. The bosonic excitations that acquire a finite energy due to long-range forces are called Higgs bosons[15] in solid-state physics, too.

Goldstone's theorem is concerned with the spectrum of excitations for systems in which a continuous symmetry of the Hamiltonian is broken. It does not tell anything about when this happens. A partial answer to this question is given by Coleman's theorem.[16] It states that quantum fluctuations will restore the continuous symmetry in the ground state of one-dimensional systems; continuous symmetry cannot be broken in the ground state, except when the order parameter associated with symmetry breaking is conserved. It was already mentioned in Chapter 2 that thermal fluctuations hinder any ordering in one-dimensional systems at finite temperatures unless the forces are long-ranged (Landau–Peierls instability). Coleman's theorem states that even the ground state cannot be ordered, since quantum fluctuations play a similar role there. The instability against breaking a continuous symmetry in two-dimensional systems at finite temperatures will be discussed in Chapters 12 and 15.

Further Reading

1. J. P. Elliott and P. G. Dawber, *Symmetry in Physics*, Vol. 1., *Principles and Simple Applications*, Macmillan Education Ltd., (1989).
2. R. A. Evarestov and V. P. Smirnov, *Site Symmetry in Crystals, Theory and Applications*, Second edition, Springer Series in Solid-State Sciences, Vol. 108, Springer-Verlag, Berlin (1997).
3. V. Heine, *Group Theory in Quantum Mechanics*, Pergamon Press, New York (1960).
4. L. D. Landau and E. M. Lifshitz, *Statistical Physics (Course of Theoretical Physics Volume 5)*, Third Edition, Butterworth-Heinemann, (1984).
5. W. Ludwig and C. Falter, *Symmetries in Physics, Group Theory Applied to Physical Problems*, Second Extended Edition, Springer-Series in Solid-State Sciences, Springer-Verlag, New York (1996).
6. G. Ya. Lyubarskii, *The Application of Group Theory in Physics*, Pergamon Press, Oxford (1960).
7. M. Tinkham, *Group Theory and Quantum Mechanics*, McGraw–Hill, New York (1964).

[15] P. W. Higgs, 1964.
[16] S. Coleman, 1973.

7

The Structure of Crystals

As we shall illustrate with several examples below, the knowledge of the space group (the geometrical symmetries of the crystal) is not sufficient to reconstruct the structure of the crystal – i.e., the position of the atoms in the primitive cell – unambiguously. Different ways of decorating the primitive cell may lead to the same space group. This means that the number of structure types is expected to be much higher than the number of space groups.

The structure of naturally occurring crystals shows a remarkable diversity indeed. However, possible structures do not appear equally frequently. The stable crystal structure of the elements at low temperature (and also at room temperature, whenever the two are different) and atmospheric pressure are listed in Appendix B. For several elements, further so-called allotropic modifications are possible. A quick glance at the table shows that about one-third of the elements are crystallized in one of four structures: cI2 and cF4 of the cubic system and hP2 and hP4 of the hexagonal system. (The nomenclature used above will be presented in the next section.) When the crystal structure of two-component compounds of composition AB is examined, the dominance of a small number of structure types is once again noticed. This is explained by the nature of the forces that hold crystals together; therefore we shall investigate the relation between crystal structure and bonding in the closing section of this chapter. Before that we shall study some frequent structures observed in elements and simple compounds – and also take a look at a handful of more complex but physically interesting structures.

7.1 Types of Crystal Structures

Two crystals are said to belong to the same structure type when their space groups and the coordinates of the atoms within the Bravais cell are identical. To be more precise, in addition to the space groups, the stoichiometric compositions have to be identical as well. However, when atoms are not in

Table 7.1. The Strukturbericht symbol, name, prototype, and Pearson symbol for a number of simple crystal structures and the Schoenflies and international notations for their space groups

Crystal structure		Proto-	Pearson	Symbols for	
Symbol	Name	type	symbol	the space group	
A1	face-centered cubic	Cu	cF4	O_h^5	$Fm\bar{3}m$
A2	body-centered cubic	W	cI2	O_h^9	$Im\bar{3}m$
A3	hexagonal close packed	Mg	hP2	D_{6h}^4	$P6_3/mmc$
A3′	double hexagonal close packed	La	hP4	D_{6h}^4	$P6_3/mmc$
A4	diamond	C	cF8	O_h^7	$Fd\bar{3}m$
A8	γ-selenium	Se	hP3	D_3^4	$P3_121$
A9	graphite	C	hP4	D_{6h}^4	$P6_3/mmc$
A15		Cr_3Si	cP8	O_h^3	$Pm\bar{3}n$
A_h	simple cubic	α-Po	cP1	O_h^1	$Pm\bar{3}m$
B1	sodium chloride	NaCl	cF8	O_h^5	$Fm\bar{3}m$
B2	cesium chloride	CsCl	cP2	O_h^1	$Pm\bar{3}m$
B3	sphalerite	ZnS	cF8	T_d^2	$F\bar{4}3m$
B4	wurtzite	ZnS	hP4	C_{6v}^4	$P6_3mc$
$B8_1$		NiAs	hP4	D_{6h}^4	$P6_3/mmc$
C1	fluorite	CaF_2	cF12	O_h^5	$Fm\bar{3}m$
C2	pyrite	FeS_2	cP12	T_h^6	$Pa\bar{3}$
C3	cuprite	Cu_2O	cP6	O_h^4	$Pn\bar{3}m$
C4	rutile	TiO_2	tP6	D_{4h}^{14}	$P4_2/mnm$
C14	Laves phase	$MgZn_2$	hP12	D_{6h}^4	$P6_3/mmc$
C15	Laves phase	Cu_2Mg	cF24	O_h^7	$Fd\bar{3}m$
$D0_2$	skutterudite	As_3Co	cI32	T_h^5	$Im\bar{3}$
$D0_3$		BiF_3	cF16	O_h^5	$Fm\bar{3}m$
$D0_9$		ReO_3	cP4	O_h^1	$Pm\bar{3}m$
$D2_f$		UB_{12}	cF52	O_h^5	$Fm\bar{3}m$
$D8_1$		Fe_3Zn_{10}	cI52	O_h^9	$Im\bar{3}m$
$E2_1$	perovskite	$CaTiO_3$	cP5	O_h^1	$Pm\bar{3}m$
$H1_1$	spinel	$MgAl_2O_4$	cF56	O_h^7	$Fd\bar{3}m$
$L1_0$		AuCu	tP2	D_{4h}^1	$P4/mmm$
$L1_2$		Cu_3Au	cP4	O_h^1	$Pm\bar{3}m$
$L2_1$	Heusler phase	$AlCu_2Mn$	cF16	O_h^5	$Fm\bar{3}m$
$L2_2$		Sb_2Tl_7	cI54	O_h^9	$Im\bar{3}m$
L′3		Fe_2N	hP3	D_{6h}^4	$P6_3/mmc$

high-symmetry positions, their coordinates may differ slightly – but the local symmetry around the atoms should be the same everywhere.

Some characteristics of the crystal structure – namely the Bravais-lattice type and the number of atoms in the Bravais cell – are contained in the *Pearson symbol*. The generic code is xYn; x is the symbol of the crystal system, given in Tables 5.7 and 5.8 (a for triclinic, m for monoclinic, o for orthorhombic, t for tetragonal, h for trigonal and hexagonal, c for cubic); Y refers to the type of the Bravais lattice (P for primitive, C for base-centered, I for body-centered, F for face-centered, R for rhombohedral); and n gives the number of atoms in the Bravais cell.

However, the Pearson symbol does not specify how the atoms are arranged within the cell. Although in many cases symmetry requirements impose restrictions on the positions of the second, third, fourth, etc. atoms (if they were placed arbitrarily, some symmetries would be broken), the Pearson symbol does not identify the structure unambiguously.

Various nomenclatures are used for specifying the structure type more precisely in the physical, chemical, and crystallographic literature. One of them is the *Strukturbericht* designation.[1] Here letters A, B, C indicate whether it is a one-component crystal, or a two-component one with a composition XY, XY_2, X_mY_n, etc. Structures within the same group are distinguished by numbers. It is also common practice to denote the structure type by the chemical symbol for a characteristic example, a prototype. Several thousand structure types exist; a handful of the simplest are selected in Table 7.1, with their usual symbols and names as well as Pearson symbols and space-group notations. The majority of the listed structures belong to the cubic system, and a smaller part to the hexagonal structure; these are the commonest for chemical elements and compounds alike. In the following sections we shall examine some of them in detail – bearing, nevertheless, in mind, that only a very narrow selection of the wealth of structures can be presented here.

7.2 Cubic Crystal Structures

Crystals that belong to the cubic crystal system come in three different Bravais-lattice types: simple, body-centered, and face-centered.

7.2.1 Simple Cubic Structures

While it is rare among the elements and not too frequent among compounds, either, our discussion starts with simple cubic crystals. This choice is justified by common practice: when some physical quantities of a crystalline material need to be determined theoretically, the lattice is most often assumed to be simple cubic.

[1] From the title of the supplement of a renowned German journal on crystallography, *Zeitschrift für Kristallographie. Strukturbericht* means *Structure Report*.

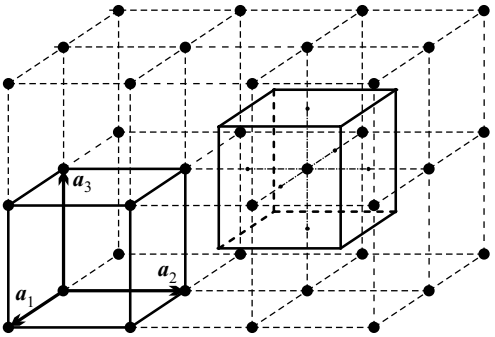


Fig. 7.1. Simple cubic lattice with the primitive vectors. The left cube (drawn with solid lines) is the primitive cell; the right one is the Wigner–Seitz cell around a lattice point

The primitive vectors of the simple cubic lattice are mutually perpendicular and of equal length. Figure 7.1 shows such a lattice with its primitive cell. Choosing a Cartesian coordinate system spanned by the unit vectors $\hat{\boldsymbol{x}}$, $\hat{\boldsymbol{y}}$, and $\hat{\boldsymbol{z}}$ along the axes of the cubic lattice, the primitive translation vectors are

$$\boldsymbol{a}_1 = a\hat{\boldsymbol{x}}\,, \qquad \boldsymbol{a}_2 = a\hat{\boldsymbol{y}}\,, \qquad \boldsymbol{a}_3 = a\hat{\boldsymbol{z}}\,, \tag{7.2.1}$$

where the edge length a of the primitive cell is the lattice constant. The volume of the primitive cell is thus $v = a^3$. The Wigner–Seitz cell is also cubic, with the same dimensions.

To determine the reciprocal lattice of the simple cubic lattice, consider matrix A_{sc} constructed from the above primitive vectors according to (5.2.3), and its inverse:

$$A_{\mathrm{sc}} = a \begin{pmatrix} 1\,0\,0 \\ 0\,1\,0 \\ 0\,0\,1 \end{pmatrix}, \qquad B_{\mathrm{sc}} = \frac{2\pi}{a} \begin{pmatrix} 1\,0\,0 \\ 0\,1\,0 \\ 0\,0\,1 \end{pmatrix}. \tag{7.2.2}$$

Using the definition (5.2.10), the primitive vectors of the reciprocal lattice are

$$\boldsymbol{b}_1 = \frac{2\pi}{a}\hat{\boldsymbol{x}}\,, \qquad \boldsymbol{b}_2 = \frac{2\pi}{a}\hat{\boldsymbol{y}}\,, \qquad \boldsymbol{b}_3 = \frac{2\pi}{a}\hat{\boldsymbol{z}}\,. \tag{7.2.3}$$

These primitive vectors generate a simple cubic lattice with lattice constant $2\pi/a$, thus the reciprocal lattice of a simple cubic lattice is another simple cubic lattice. Choosing the Brillouin zone as the domain of inequivalent $\boldsymbol{k}$ vectors instead of the primitive cell spanned by the $\boldsymbol{b}_i$, this is also found to be cubic, with the same edge length, $2\pi/a$. The volume of the Brillouin zone is thus $(2\pi)^3/a^3$.

In the previous chapter it was shown that the eigenstates of a crystalline system may be specified in terms of the wave vectors $\boldsymbol{k}$ (as quantum numbers) defined in the Brillouin zone. We shall extensively use the convention of

denoting special points of high symmetry by specific letters in the description of lattice vibrations and electronic states. The special points of the Brillouin zone of a simple cubic lattice and their letter symbols are shown in Fig. 7.2.

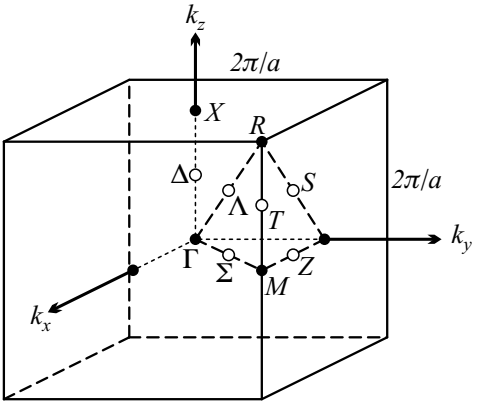


Fig. 7.2. The Brillouin zone of a simple cubic lattice with the special points and their letter symbols

The point group of a simple cubic (cP) Bravais lattice is $m\bar{3}m$ (O_h). We have seen that, depending on the choice of the basis, a part of the symmetries may be broken. Below we shall mention some structures in which the introduction of a basis does not break any of the symmetries of the cube – that is, whose space group is $Pm\bar{3}m$ (O_h^1). This imposes severe restrictions on atomic positions. The structures will be specified by the number and position of the atoms in the basis.

The simplest crystal has one atom per unit cell; its Pearson symbol is cP1, and its Strukturbericht symbol is A_h. The unit cell contains one atom, and so lattice points can be chosen as atomic positions. This is shown in Fig. 7.3(a). Very few natural crystals have this structure: among the elements α-polonium is the only one.

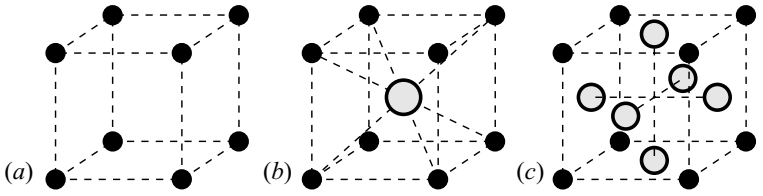


Fig. 7.3. Primitive cell for simple cubic structures with one-, two-, and four-point bases. (a) Monatomic cP1 (A_h or Po) structure; (b) diatomic cP2 (B2 or CsCl) structure; (c) tetratomic cP4 ($L1_2$ or Cu_3Au) structure

The reason for this is the bad space filling of this arrangement: each atom has only six nearest neighbors. The number of nearest neighbors, called the *coordination number*, is an important property of the crystal. In a monatomic simple cubic structure this number is thus 6. Even though atoms are in contact along the axes, large volumes are left empty among them in the interior of the cells. Space filling can be quantitatively characterized by the *packing fraction* (also called the *atomic packing factor*, APF), the ratio of the volume of the atoms considered as touching spheres to the volume of the cell. When a monatomic simple cubic lattice is constructed from spheres of radius r in contact with each other, the edge length is $a = 2r$. The volume occupied by the (single) atom is $4r^3\pi/3 = \pi a^3/6$. This gives a packing fraction of $\pi/6 = 0.524$.

The lattice is more densely filled when the empty region among the atoms sitting at the vertices of the cubes is occupied by further atoms. When the basis is made up of two atoms, cubic symmetry is preserved only if one of them is in a vertex and the other is at the center of the primitive cell. In this arrangement – shown in Fig. 7.3(b) – the primitive cell contains two atoms, hence the Pearson symbol is cP2.

A typical example for simple cubic structures with such a basis is cesium chloride (CsCl), thus this arrangement is frequently called *CsCl structure*. The traditional crystallographic designation is *B2 structure*. Positively charged Cs^+ ions (cations) sit in point 000 of each primitive cell, while negatively charged Cl^- ions (anions) sit at the centers $\frac{1}{2}\frac{1}{2}\frac{1}{2}$. It is readily seen in the figure that each ion is surrounded by eight oppositely charged ions forming a cubic environment, the coordination number is therefore 8. Among two-component compounds of composition AB this structure ranks among the most frequently occurring ones. Among others, brass (CuZn, β-brass) belongs to this group.

A cubic lattice cannot be decorated with a three-point basis in such a way that cubic symmetry is preserved. However, when the basis contains four atoms, several possibilities exist. In one of them, shown in Fig. 7.3(c), atoms are found at the vertices and face centers, and the chemical composition is AB_3. This structure is denoted by cP4 or (in traditional crystallographic notation) by $L1_2$. A typical example is the ordered phase of the alloy Cu_3Au, which is why it is also called Cu_3Au structure. The "A" atoms (in the above example, Au) sit at the vertices, while the "B" atoms (Cu) sit at the face centers. Thus the atomic coordinates in the basis are 000 for the Au atom and $0\frac{1}{2}\frac{1}{2}$, $\frac{1}{2}0\frac{1}{2}$, and $\frac{1}{2}\frac{1}{2}0$ for the Cu atoms. Each atom has 12 nearest neighbors in this array, the coordination number is therefore 12. Au atoms are surrounded by Cu atoms exclusively, and cubic (O_h) symmetry is locally preserved. Among the nearest neighbors of Cu atoms there are four Au and eight Cu atoms; the local symmetry is only tetragonal (D_{4h}).

Some further examples with more complicated bases are shown in Fig. 7.4; symmetries of the space group $Pm\bar{3}m$ (O_h^1) are preserved in each arrangement. The primitive cell contains four atoms in the ReO_3 (or $D0_9$) structure, too,

thus its Pearson symbol is also cP4. Rhenium atoms sit at the vertices and oxygen atoms at the midpoints of the edges, i.e., at $\frac{1}{2}00$, $0\frac{1}{2}0$, $00\frac{1}{2}$.

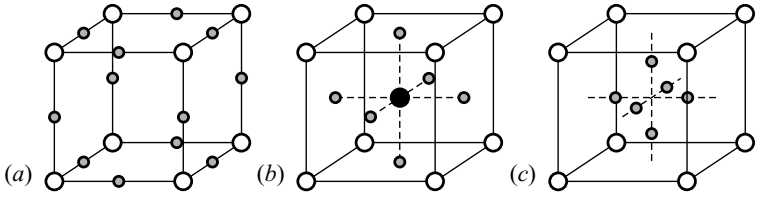


Fig. 7.4. Simple cubic structures with four, five, and seven atoms per primitive cell. (*a*) cP4 ($D0_9$); (*b*) cP5 ($E2_1$ or perovskite); (*c*) cP7 ($D2_1$) structure

The cubic perovskite structure[2] contains five atoms per primitive cell. Its traditional symbol is $E2_1$ and its Pearson symbol is cP5. In its prototype, $CaTiO_3$, the large cations (Ca^{2+}) sit at the vertices of the cube, the small cations (Ti^{4+}) at the cell center, and the large anions (O^{2-}) at the face centers $0\frac{1}{2}\frac{1}{2}$, $\frac{1}{2}0\frac{1}{2}$, $\frac{1}{2}\frac{1}{2}0$. The large cation is thus surrounded by 12 anions, while the small cation by just 6.

The basis is composed of seven atoms in the CaB_6 structure ($D2_1$, cP7). Here Ca ions sit at the vertices of the cube, and the six boron ions inside the primitive cell are along the lines between the body center and the face centers. The distance from the body center is such that each boron atom is surrounded by five others at equal distances – at points that are not of particularly high symmetry otherwise. The coordinates of the boron ions are $(\frac{1}{2} \pm x)\frac{1}{2}\frac{1}{2}$; $\frac{1}{2}(\frac{1}{2} \pm x)\frac{1}{2}$, and $\frac{1}{2}\frac{1}{2}(\frac{1}{2} \pm x)$, where $x = 1 - \sqrt{2}/2 = 0.293$. Their coordination number is 5 – while that of Ca ions is much higher, 24.

Other choices for the basis may also preserve the full symmetry of the space group $Pm\bar{3}m$. For example, when the composition of the compound/alloy is AB, with one type of atom at the face centers of the cell and the other at the edge centers. This structure (denoted by the Pearson symbol cP6) is observed in NbO. Symmetries remain the same when a third kind of atom, C is placed at the vertices (A_3B_3C). Much more complicated is the $BaHg_{11}$ ($D2_e$, cP36) structure. Here barium atoms sit at the edge centers, one mercury atom is located at the center of the cell, and the others – 32 further mercury atoms per primitive cell – around the vertices and edge centers in an array that preserves symmetry.

Naturally, in many more cases the symmetry of the crystal is reduced because of the lower symmetry of the basis. We shall briefly mention just two of them, the A15 and C3 structures. The first one is of particular importance as materials with this structure held the record for the superconducting transition temperature for many years, and this was essentially due to some

[2] After the mineralogical name of $CaTiO_3$ – although the actual structure of this mineral is distorted in its ground state: its space group is only *Pnma*.

characteristics of the structure. The primitive cell of the A15[3] (Cr_3Si) structure contains the atoms of two molecules; its Pearson symbol is therefore cP8. As it is illustrated in Fig. 7.5, silicon atoms are located at the vertices and body centers of the cells, while chromium atoms are found at points of lower symmetry, $\frac{1}{2}0\frac{1}{4}$ and $\frac{1}{2}0\frac{3}{4}$ – as well as $0\frac{1}{4}\frac{1}{2}$, $0\frac{3}{4}\frac{1}{2}$, $\frac{1}{4}\frac{1}{2}0$, and $\frac{3}{4}\frac{1}{2}0$, which are obtained from the former two by rotations around the space diagonal. Each Si atom is thus surrounded by 12 Cr atoms in an icosahedral arrangement. Fourfold rotations around the x-, y-, and z-axes are no longer symmetries, unless they are followed by a translation along the face diagonal perpendicular to the axis of rotation. Rotation axes are thus replaced by screw axes. The space group of A15 compounds is therefore $Pm\bar{3}n$ (O_h^3).

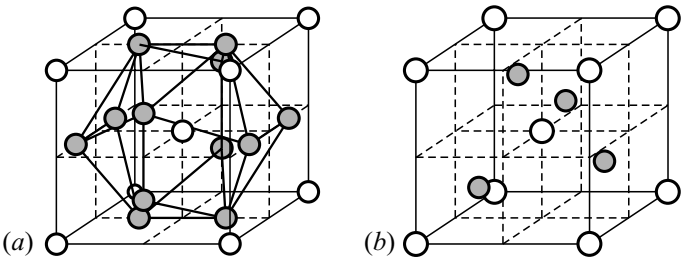


Fig. 7.5. Atomic arrangement in the primitive cell of compounds (a) of composition AB_3 and structure A15; (b) of composition A_2B and structure C3

The prototype of the C3 structure is cuprite, Cu_2O. The primitive cell contains two molecules, whereby its Pearson symbol is cP6. More generally, the composition of the material is A_2B; atoms of type B are now located at the vertices and body centers of each primitive cell, while atoms of type A are found at $\frac{1}{4}\frac{1}{4}\frac{1}{4}$, $\frac{3}{4}\frac{3}{4}\frac{1}{4}$, $\frac{3}{4}\frac{1}{4}\frac{3}{4}$, and $\frac{1}{4}\frac{3}{4}\frac{3}{4}$. Once again, fourfold rotations are no longer symmetries unless combined with a translation along the space diagonal. Moreover, the same applies to inversion in this case. The space group is therefore $Pn\bar{3}m$ (O_h^4).

7.2.2 Body-Centered Cubic Structures

In the structures presented in the previous subsection either the atoms located at the vertices were different from those at the body- or face-centers of the cubic cell, or the atoms were of the same species but the basis was such that the surroundings of the atoms at the vertices and at the centers were different. In either case, the edge vectors of the cube were the primitive vectors. A different

[3] Recall that in the Strukturbericht designation the letter A is reserved for structures that occur in chemical elements. An elemental example for the A15 structure is the metastable β phase of tungsten (W).

situation arises when the atoms or groups of atoms sitting at the vertices and at the centers of the cubic cells are the same and so are their surroundings. Then the cubic unit cell is not a primitive cell any more, as the crystal is taken into itself by a translation through half the space diagonal of the cell. The Bravais lattice is thus extended to include body centers in addition to vertices of the cube. The new lattice is called *body-centered cubic* (bcc).

The body-centered cubic lattice can be conceived to be built up of two interpenetrating simple cubic lattices that are displaced by half the space diagonal of the cubic cell with respect to one another. It is then customary to speak of two sublattices. It should be emphasized that if the two sublattices become inequivalent – for example, when the two components of a disordered alloy start to become ordered on separate sublattices in a disorder–order transition – then the initial structure, a body-centered cubic lattice with a monatomic basis is transformed into a simple cubic lattice with a diatomic basis.

Primitive vectors of the bcc lattice are usually chosen to join the center of the cube with three vertices, as illustrated in Fig. 7.6(a). Denoting the edge length of the cube by a,

$$\boldsymbol{a}_1 = \frac{a}{2}(-\hat{\boldsymbol{x}} + \hat{\boldsymbol{y}} + \hat{\boldsymbol{z}})\,, \quad \boldsymbol{a}_2 = \frac{a}{2}(\hat{\boldsymbol{x}} - \hat{\boldsymbol{y}} + \hat{\boldsymbol{z}})\,, \quad \boldsymbol{a}_3 = \frac{a}{2}(\hat{\boldsymbol{x}} + \hat{\boldsymbol{y}} - \hat{\boldsymbol{z}})\,. \tag{7.2.4}$$

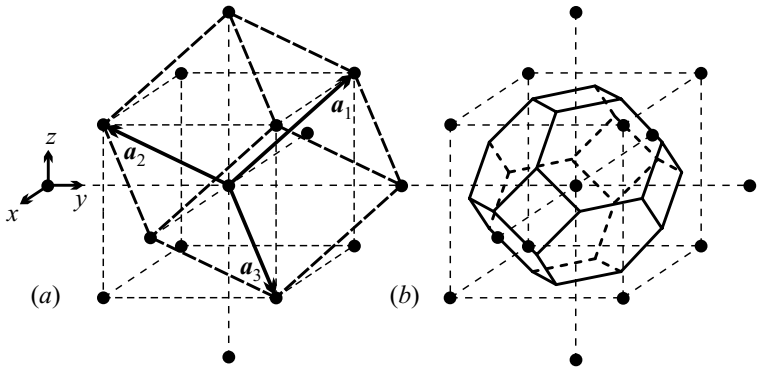


Fig. 7.6. Body-centered cubic crystals: (a) the primitive cell with the primitive vectors; (b) the Wigner–Seitz cell

The figure clearly shows that the primitive cell of volume $a^3/2$ – spanned by the primitive translation vectors – is rhombohedral, therefore it does not possess the symmetries of a cube. For this reason, the cubic Bravais cell (conventional unit cell) is preferred. Its edge vectors are given by

$$\boldsymbol{a} = \boldsymbol{a}_2 + \boldsymbol{a}_3 = a\hat{\boldsymbol{x}}\,, \quad \boldsymbol{b} = \boldsymbol{a}_1 + \boldsymbol{a}_3 = a\hat{\boldsymbol{y}}\,, \quad \boldsymbol{c} = \boldsymbol{a}_1 + \boldsymbol{a}_2 = a\hat{\boldsymbol{z}}\,. \tag{7.2.5}$$

Instead of the primitive cell Fig. 7.6(a) the Wigner–Seitz cell is sometimes used; their volumes are equal but the latter is more symmetric. To

construct the Wigner–Seitz cell consider the body center of a Bravais cell. It has eight nearest neighbor lattice sites in space-diagonal directions and six second neighbor sites, located at the centers of adjacent cubes. The object bounded by the perpendicular bisector planes of the segments between the body center and these neighbors is an Archimedean solid, a truncated regular octahedron. In the directions of the eight nearest neighbors its faces are regular hexagons, while in the directions of the six second neighbors they are squares. The Wigner–Seitz cell, shown in Fig. 7.6(b), possesses every symmetry of the cube.

To specify the reciprocal lattice, first the matrix A is expressed in terms of the reciprocal vectors – see (5.2.3) –, and then its inverse is taken:

$$\mathsf{A}_{\mathrm{bcc}} = \frac{a}{2}\begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix}, \qquad \mathsf{B}_{\mathrm{bcc}} = \frac{2\pi}{a}\begin{pmatrix} 0\,1\,1 \\ 1\,0\,1 \\ 1\,1\,0 \end{pmatrix}. \tag{7.2.6}$$

Using (5.2.10), the primitive vectors of the reciprocal lattice are

$$\boldsymbol{b}_1 = \frac{2\pi}{a}(\hat{\boldsymbol{y}} + \hat{\boldsymbol{z}})\,, \qquad \boldsymbol{b}_2 = \frac{2\pi}{a}(\hat{\boldsymbol{z}} + \hat{\boldsymbol{x}})\,, \qquad \boldsymbol{b}_3 = \frac{2\pi}{a}(\hat{\boldsymbol{x}} + \hat{\boldsymbol{y}})\,. \tag{7.2.7}$$

The reciprocal lattice generated by these vectors is a face-centered cubic lattice of edge $4\pi/a$, as shown in Fig. 7.7. The Brillouin zone – obtained via Dirichlet's construction – is drawn with dashed lines in the left part; special points of the Brillouin zone and their letter symbols are marked in the right part.

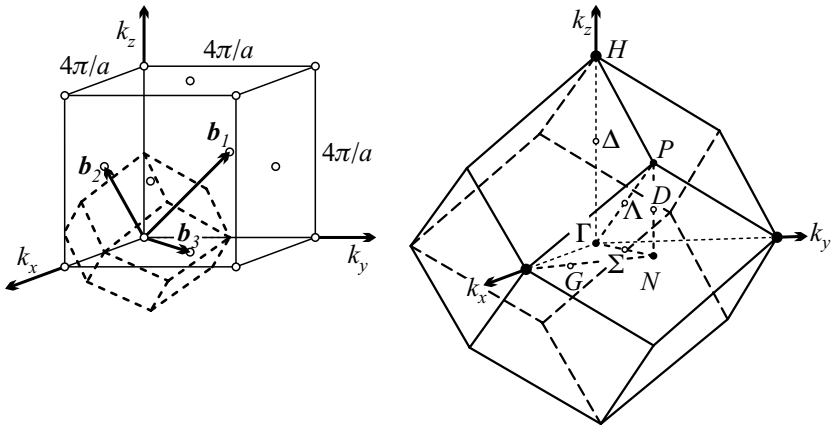


Fig. 7.7. The reciprocal lattice of a body-centered cubic lattice and its Brillouin zone with special points

The Pearson symbol for a body-centered cubic structure, with a single atom per lattice point in the Bravais lattice is cI2, as each Bravais cell contains

two atoms. As listed in the table of Appendix B, a relatively large number of elements crystallize in this structure at room temperature, e.g., monovalent alkali metals (Li, Na, K, Rb, Cs), two heavy alkaline-earth metals (Ba, Ra), a transition metal (Cr), two forms (α and δ) of iron (Fe), and also Mo, Nb, Ta, V, and W. The latter one can be considered as the prototype, hence its traditional name: *W structure* (in the Strukturbericht notation: A2 structure).

This structure occurs more frequently than the simple cubic structure among the elements. This is because space filling is more efficient (but still not ideal) with an extra atom at the center of the cube. Each atom is surrounded by eight others in a cubic arrangement, the coordination number is thus 8. When the lattice is filled with one kind of spherical atom of radius r in such a way that atoms at the vertices touch those at the body centers, the space diagonal of the Bravais lattice is $4r$ – that is, the edge length of the cubic Bravais cell is $a = 4r/\sqrt{3}$, and so its volume is $(4r/\sqrt{3})^3$. Since each Bravais cell contains two atoms, the packing fraction is

$$2\frac{4}{3}r^3\pi\left(\frac{4r}{\sqrt{3}}\right)^{-3} = \frac{\sqrt{3}}{8}\pi = 0.680\,. \tag{7.2.8}$$

The W structure, which has a monatomic basis, possesses all rotation and reflection symmetries of a cube; its space group is $Im\bar{3}m$ (O_h^9). When the basis consists of several atoms, there exist a number of structures for which the same space group is preserved. In La_2O_3 lanthanum atoms are located at vertices and body centers, and oxygen atoms at face and edge centers, however, on the average only half of the possible oxygen positions are occupied. The Pearson symbol for this arrangement is thus cI5. These positions or sites – face and edge centers – are of particular importance, as they are highly symmetric, surrounded in an octahedral arrangement by the atoms at the vertices and body centers of the Bravais lattice, as illustrated in Fig. 7.8(a). Besides these *octahedral sites* the arrangement has other high-symmetry positions: $(\frac{1}{4}\frac{1}{2}0)$ and all the sites obtained from this through the symmetry transformations of the cube. These are called *tetrahedral sites* because of the tetrahedral arrangement of the nearest neighbors (Fig. 7.8(b)). Note that the A15 structure shown in Fig. 7.5(a) can also be regarded as an arrangement in which not only the vertices and centers of the Bravais lattice of a body-centered crystal are occupied but also half of the tetrahedral sites – in such a manner that the space group is $Pm\bar{3}n$.

There exist more complicated structures in which the symmetry of the space group $Im\bar{3}m$ is preserved, e.g., $D8_1$, $D8_f$, and $L2_2$ structures, whose prototypes are Fe_3Zn_{10}, Ir_3Ge_7, and Sb_2Tl_7. In the first two cases the primitive cell contains two molecules and the Bravais cell four, while in the third case there are three molecules per primitive cell and six per Bravais cell. Their Pearson symbols are therefore cI52, cI40, and cI54.

In simpler structures some of the symmetries are broken because of the lower symmetry of the basis. Of particular interest is the $MoAl_{12}$ structure,

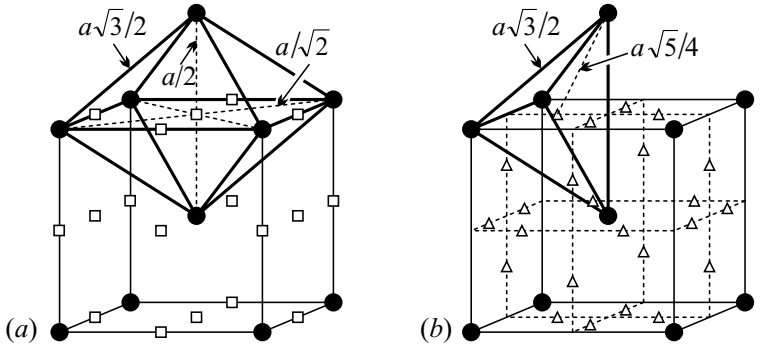


Fig. 7.8. Highly symmetric empty sites in a bcc lattice: (*a*) octahedral sites (□); (*b*) tetrahedral sites (△)

illustrated in Fig. 7.9. Molybdenum atoms are located at the vertices and centers of the cubic Bravais cells. Each of them are surrounded by twelve aluminum atoms in an icosahedral arrangement. The space group is therefore reduced to $Im\bar{3}$ (T_h^5). Each Bravais cell contains two units, hence the Pearson symbol is cI26. The compounds $MnAl_{12}$ and $Al_{12}W$ both crystallize in this structure; the latter is considered as the prototype.

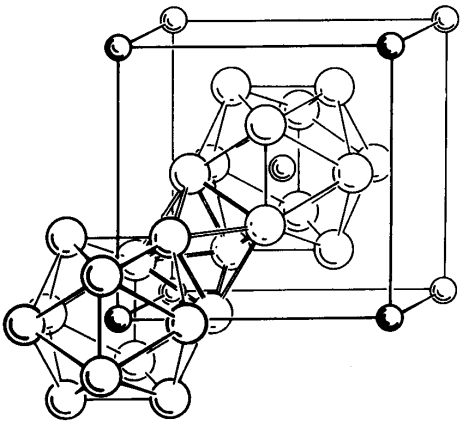


Fig. 7.9. $MoAl_{12}$ structure. For the sake of clarity, the twelve icosahedrally arranged Al atoms are drawn only around two sites, a vertex and a body center

7.2.3 Face-Centered Cubic Structures

Face-centered cubic (fcc) crystals are obtained by decorating an fcc Bravais lattice with identical groups of atoms. Disordered alloys (e.g., Cu_3Au), in

which the two components are found at each lattice point with equal probability also belong here. Upon heating, the simple cubic structure of Cu_3Au (space group: $Pm\overline{3}m$) is transformed into a face-centered one at a critical temperature.

Primitive vectors point from the vertices of the Bravais cell into the face centers:

$$\boldsymbol{a}_1 = \frac{a}{2}(\hat{\boldsymbol{y}} + \hat{\boldsymbol{z}})\,, \qquad \boldsymbol{a}_2 = \frac{a}{2}(\hat{\boldsymbol{x}} + \hat{\boldsymbol{z}})\,, \qquad \boldsymbol{a}_3 = \frac{a}{2}(\hat{\boldsymbol{x}} + \hat{\boldsymbol{y}})\,. \tag{7.2.9}$$

The face-centered cubic lattice can therefore be viewed as four interpenetrating simple cubic lattices, whose origins are displaced by the above vectors. The lattice is thus decomposed into four equivalent sublattices. The primitive cell spanned by the primitive vectors is rhombohedral, as shown in Fig. 7.10(a), so it does not possess the symmetries of a cube. This is why preference is often given to the conventional unit cell, which is four times bigger but more symmetric: the cubic Bravais cell with edge vectors

$$\begin{aligned} \boldsymbol{a} &= -\boldsymbol{a}_1 + \boldsymbol{a}_2 + \boldsymbol{a}_3 = a\hat{\boldsymbol{x}}\,, \\ \boldsymbol{b} &= \boldsymbol{a}_1 - \boldsymbol{a}_2 + \boldsymbol{a}_3 = a\hat{\boldsymbol{y}}\,, \\ \boldsymbol{c} &= \boldsymbol{a}_1 + \boldsymbol{a}_2 - \boldsymbol{a}_3 = a\hat{\boldsymbol{z}}\,. \end{aligned} \tag{7.2.10}$$

In what follows, we shall specify atomic positions in this Bravais cell.

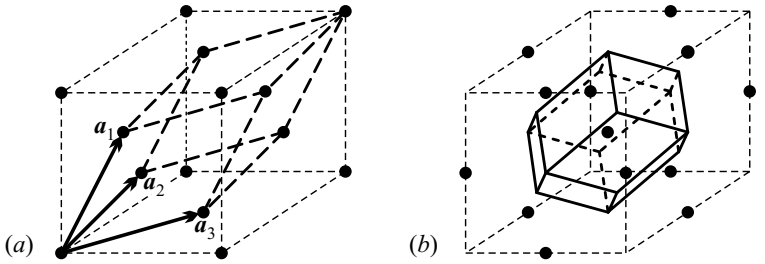


Fig. 7.10. Face-centered cubic crystals: (a) the Bravais cell with the primitive vectors; (b) the Wigner–Seitz cell

The symmetric Wigner–Seitz cell, obtained via Dirichlet's construction, is shown in Fig. 7.10(b). For clarity, it is drawn into a cube that is displaced by half the space diagonal with respect to the Bravais cell – that is, lattice points are at the body and edge centers of the cube. Since each lattice point has 12 nearest neighbors, when the cube is cut by the perpendicular planes through the midpoints, a rhombic dodecahedron is obtained for the Wigner–Seitz cell. In the direction of the 12 adjacent lattice sites it is bordered by congruent rhombi. The Wigner–Seitz cell shows every symmetry of the cube. As we have seen, the Brillouin zone of a body-centered cubic crystal has the same shape.

The reciprocal lattice of an fcc lattice is determined in the usual way, by taking the inverse of matrix A constructed from the primitive vectors of the direct lattice:

$$A_{\mathrm{fcc}} = \frac{a}{2}\begin{pmatrix} 0\,1\,1 \\ 1\,0\,1 \\ 1\,1\,0 \end{pmatrix}, \qquad B_{\mathrm{fcc}} = \frac{2\pi}{a}\begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix}. \tag{7.2.11}$$

The reciprocal-lattice vectors are then

$$\begin{aligned} \boldsymbol{b}_1 &= \frac{2\pi}{a}(-1,1,1) = \frac{2\pi}{a}(-\hat{\boldsymbol{x}} + \hat{\boldsymbol{y}} + \hat{\boldsymbol{z}})\,, \\ \boldsymbol{b}_2 &= \frac{2\pi}{a}(1,-1,1) = \frac{2\pi}{a}(\hat{\boldsymbol{x}} - \hat{\boldsymbol{y}} + \hat{\boldsymbol{z}})\,, \\ \boldsymbol{b}_3 &= \frac{2\pi}{a}(1,1,-1) = \frac{2\pi}{a}(\hat{\boldsymbol{x}} + \hat{\boldsymbol{y}} - \hat{\boldsymbol{z}})\,. \end{aligned} \tag{7.2.12}$$

Comparison with (7.2.4) shows that the reciprocal of a face-centered cubic lattice is a body-centered cubic lattice with a lattice constant of $4\pi/a$. Face- and body-centered cubic lattices are thus reciprocal to one another in the sense that the reciprocal lattice of an fcc lattice is a bcc lattice, and vice versa. This implies that the Brillouin zone of an fcc lattice is of the same shape as the Wigner–Seitz cell of a bcc lattice: a truncated octahedron. This is shown in Fig. 7.11, along with the special points of the Brillouin zone.

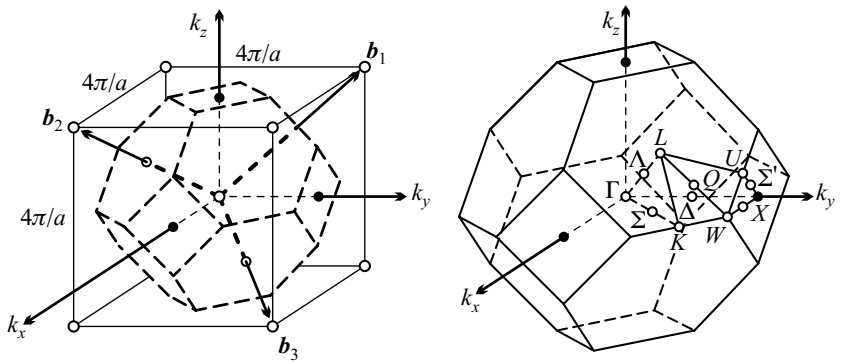


Fig. 7.11. The reciprocal lattice of the face-centered cubic lattice and its Brillouin zone with special points

The simplest crystal structure in a face-centered (F) cubic (c) lattice with a monatomic basis contains four atoms per Bravais cell, since only 1/8 of each atom at the vertices and 1/2 of each atom at the face centers belong to the cell. This arrangement is thus denoted by cF4. Each atom is surrounded by twelve identical ones as nearest neighbors, so the coordination number is 12. For spherical atoms with equal radii this is the highest possible number of

touching neighbors. When the radius of the spheres is r, it is related to the edge length of the cube by $4r = \sqrt{2}a$. Since the Bravais cell of volume

$$\left(\frac{4}{\sqrt{2}}r\right)^3 = \frac{32}{\sqrt{2}}r^3 \tag{7.2.13}$$

contains four atoms, the packing fraction for this crystal structure is

$$4\,\frac{4}{3}r^3\pi\left(\frac{32}{\sqrt{2}}r^3\right)^{-1} = \frac{\sqrt{2}}{6}\pi = 0.740\,. \tag{7.2.14}$$

This is higher than the values obtained for simple and body-centered cubic crystals. Indeed, spherical atoms fill the space optimally in this arrangement. This is why the face-centered cubic structure is also known as the *cubic close-packed* (ccp) structure.

Close packing is best illustrated by the atomic arrangement in the (111) atomic planes perpendicular to the space diagonal. As it can be seen in Fig. 7.12, within these planes atoms form a hexagonal lattice, which is the most efficient way of covering the plane with circles. In the next section we shall return to the question of how to stack such atomic layers to obtain an fcc or some other structure.

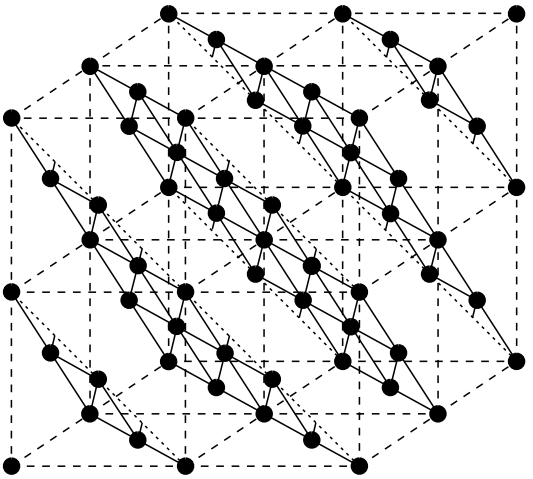


Fig. 7.12. Close packing of atoms in an fcc crystal, in the planes perpendicular to the space diagonal

As close packing is favored by the metallic bond, many metals crystallize in this structure. The prototype, after which this monatomic structure has been named, is copper. The space group of the Cu structure (also called A1 structure) is $Fm\bar{3}m$ (O_h^5). Besides copper, several other elements crystallize

in this form: other noble metals, Ag and Au; two alkaline-earth metals, Ca, Sr; trivalent Al; several transition metals such as γ-Fe, α-Co, Ni, Ir, Pt, Rh, Pd, and Pb; some lanthanoids and actinoids, e.g. Ce, Pr, Th, Yb; and also a couple of noble gases in their low-temperature solid phase, Ar, Ne, Kr, and Xe.

In addition to the vertices and face centers of the cube, atoms have to be placed at other sites as well to obtain face-centered cubic structures with a polyatomic basis. These additional atoms tend to fill up the empty spaces among the atoms at the vertices and face centers. In face-centered cubic structures there are two typical sites that can be occupied.

One of them is the center of the cube, site $\frac{1}{2}\frac{1}{2}\frac{1}{2}$, and the equivalent positions at the midpoints of the edges, which can be reached by translations of the cube center through primitive vectors. These sites are surrounded by six lattice points in an octahedral geometry, therefore they are also called *octahedral sites*. Rotations and reflections that leave such a point invariant and transform the crystal into itself are just the elements of the point group O_h.

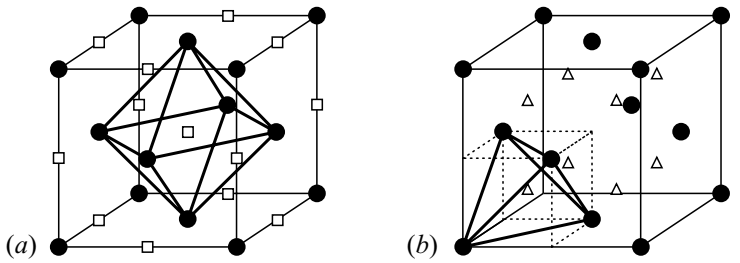


Fig. 7.13. Highly symmetric empty sites in an fcc lattice: (*a*) octahedral sites (□); (*b*) tetrahedral sites (△)

Other typical sites are the centers of the octants of the cube: $\frac{1}{4}\frac{1}{4}\frac{1}{4}$ and equivalent points. When an atom is placed there, its four nearest neighbors in the fcc lattice form a regular tetrahedron. For this reason they are called *tetrahedral sites*. The local symmetry group at these sites is T_d. The two types of site are shown in Fig. 7.13. At each site of the lattice the potential due to all other atoms (called the crystal-field potential) shows the symmetries of the point group of the site in question. As discussed in the previous chapter, this potential may give rise to crystal-field splitting. Since the main features of the splitting (apart from its magnitude) depend on the local symmetry, knowing it is important for understanding atomic energy spectra. Conversely, the type of splitting can be used to determine the local symmetry – and hence the position of the atoms.

There are eight tetrahedral and four octahedral sites in a Bravais cell. In the simplest fcc crystals with multiatomic bases, besides vertices and face centers these sites are occupied by atoms – either completely or partially. De-

pending on the occupancy of each site, various structures are possible. Because of their importance and the partial symmetry breaking occurring in them, diamond structure and the closely related sphalerite structure merit separate discussion. Below we shall present some other types in which symmetries of the space group $Fm\bar{3}m$ (O_h^5) are fully preserved.

A common structure is the *sodium chloride* or *rock-salt structure* (in traditional notation: B1 structure). Sodium cations (Na^+) are located at vertices and face centers – at 000, $0\frac{1}{2}\frac{1}{2}$, $\frac{1}{2}0\frac{1}{2}$, and $\frac{1}{2}\frac{1}{2}0$, if the edge of the fcc Bravais cell is chosen as unity –, and chlorine anions (Cl^-) at the octahedral sites $\frac{1}{2}00$, $0\frac{1}{2}0$, $00\frac{1}{2}$, and $\frac{1}{2}\frac{1}{2}\frac{1}{2}$. This arrangement is illustrated in Fig. 7.14(a).

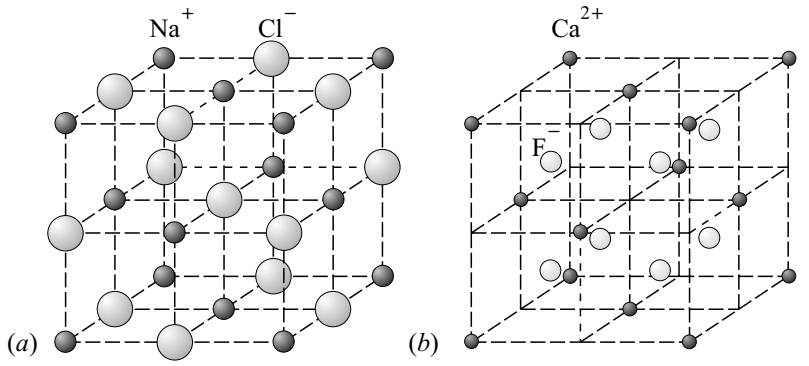


Fig. 7.14. Face-centered cubic structures: (a) sodium chloride (NaCl, B1) structure; (b) fluorite (C1) structure

Sodium and chlorine ions are at alternate points of a simple cubic lattice, forming a three-dimensional checkerboard pattern. Each Na^+ ion is surrounded by six Cl^- ions and vice versa, the coordination number is thus 6. Cations and anions form two interpenetrating fcc sublattices that are displaced through half the space diagonal with respect to one another. From the viewpoint of symmetries, the overall structure is face-centered cubic, too. Since each Bravais cell contains four sodium chloride molecules, the Pearson symbol is cF8. The four times smaller primitive cell contains only two ions, Na^+ at 000 and Cl^- at $\frac{1}{2}\frac{1}{2}\frac{1}{2}$. These two ions make up the basis of the lattice. The halides of all alkali metals but Cs crystallize in the rock-salt structure – and so do divalent salts, e.g., MgO, CaO, MgS, CaSe, and BaTe. We shall give the simple geometrical reason for this at the end of the chapter.

Fluorite structure or C1 structure – the prototype of which is fluorite (CaF_2) – is obtained by placing atoms of the second type at the tetrahedral (rather than octahedral) sites. This structure is shown in Fig. 7.14(b). Calcium cations (Ca^{2+}) are located at vertices and face centers, and fluorine anions (F^-) at the centers of the octants. Translation of the basis – made up of a Ca^{2+} ion at 000 and two F^- ions at $\frac{1}{4}\frac{1}{4}\frac{1}{4}$ and $\frac{3}{4}\frac{3}{4}\frac{3}{4}$ – through the primitive

vectors (7.2.9) generates the full crystal. The Bravais cell contains four CaF_2 molecules, so the Pearson symbol is cF12. Each Ca^{2+} ion is surrounded by eight F^- ions in a cubic arrangement, their coordination number is thus 8; F^- ions, on the other hand, are surrounded by only four Ca^{2+} ions in a tetrahedral arrangement, so their coordination number is just 4. The oxides and sulfides of alkali metals (e.g., Li_2O, Na_2O, K_2O, Na_2S, K_2S), as well as many other oxides and halides crystallize in this structure.

When both octahedral and tetrahedral sites are occupied by atoms of the same (second) type, a BiF_3 or $D0_3$ structure – also known as $AlFe_3$ structure – arises. This is shown in Fig. 7.15(a). Bismuth cations (Bi^{3+}) are located at the vertices and face centers of the cube, and fluorine anions (F^-) among them, at octahedral and tetrahedral sites. The Bravais cell contains four molecules, and so its Pearson symbol is cF16.

The octahedral and tetrahedral sites are all occupied – although by two different types of ion – in the $L2_1$ structure. The prototype for this structure is the Heusler alloy $AlMnCu_2$, which is particularly noted for its magnetic properties.

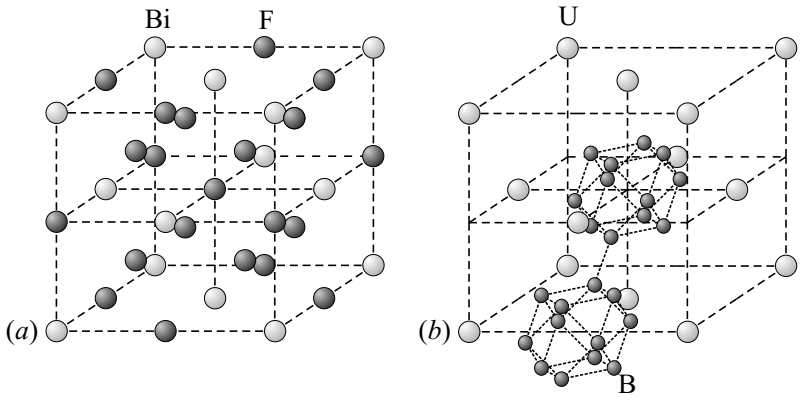


Fig. 7.15. Face-centered cubic structures: (a) BiF_3 structure ($D0_3$ structure), with 16 atoms per Bravais cell; (b) UB_{12} structure ($D2_f$ structure) with 52 atoms per Bravais cell. Only two of the four cuboctahedra are shown; each contains twelve boron atoms

Even more complicated is the $D2_f$ structure shown in Fig. 7.15(b), whose prototype is UB_{12}. Among the uranium atoms at the vertices and face centers, octahedral sites are occupied not by single ions but groups of twelve boron atoms forming a cuboctahedron.[4] Referred to the octahedral site (e.g., the center of the Bravais cell), the coordinates of the 12 boron atoms around it are $0(\pm\frac{1}{6})(\pm\frac{1}{6})$; $(\pm\frac{1}{6})0(\pm\frac{1}{6})$, and $(\pm\frac{1}{6})(\pm\frac{1}{6})0$. With this choice the distance

[4] Also called a heptaparallelohedron or triangular gyrobicupola; a cube truncated at its vertices to its edge centers.

between atoms that belong to adjacent cuboctahedra is the same as the interatomic distance within a cuboctahedron. The Bravais cell contains four molecules, hence the Pearson symbol is cF52.

One may find face-centered cubic structures in which the basis contains a much higher number of atoms. The Bravais lattice of the $NaZn_{13}$ or $D2_3$ structure contains 112 atoms. An even higher number of atoms is found in the basis of crystalline fullerite, formed by C_{60} molecules (shown in Fig. 2.8). In this allotropic form of carbon a C_{60} molecule is sitting at each point of an fcc lattice. With carbon atoms arranged in regular pentagonal and deformed hexagonal rings at the vertices of a truncated icosahedron, the molecule itself shows the symmetries of the icosahedral point group I_h ($m\bar{3}\bar{5}$).[5] If the orientation of the fivefold rotation axes of the C_{60} molecules is ordered, then – because of the incompatibility of cubic and icosahedral symmetries – the entire crystal cannot show each symmetry of the cube. The space group of fullerite is $Pa\bar{3}$ (T_h^6) in this low-temperature ordered phase. At higher temperatures, in the plastic-crystalline phase, fullerene molecules are free to rotate, and thus full cubic symmetry is restored. The space group of the crystal is then $Fm\bar{3}m$ (O_h^5).

7.2.4 Diamond and Sphalerite Structures

Among fcc structures with a diatomic basis particularly important are those in which all vertices and face centers but only half of the tetrahedral sites are occupied. Depending on whether the atoms at the tetrahedral sites are identical with those at the vertices and face centers, the arrangement is either a *diamond structure* or a *sphalerite (zincblende) structure*. Their Strukturbericht designations are A4 and B3. In either case, the basis consists of two atoms, at 000 and $\frac{1}{4}\frac{1}{4}\frac{1}{4}$. The Bravais cell contains eight atoms, thus the Pearson symbol is cF8 for both of them. The two structures are shown in Fig. 7.16.

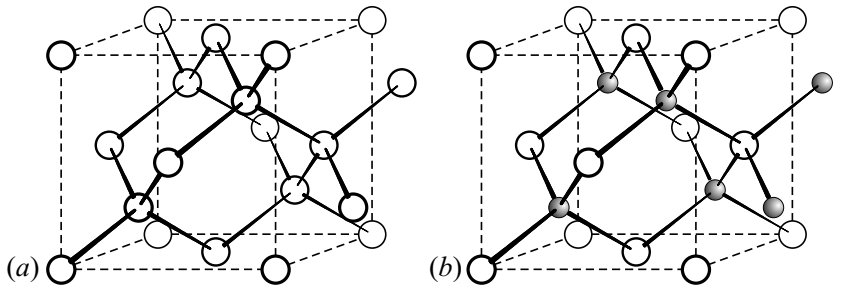


Fig. 7.16. Fcc crystal structures with a two-point basis: (*a*) diamond (A4) structure; (*b*) sphalerite (B3) structure

[5] Pentagons are bounded by hexagons, while hexagons are bounded by pentagons and hexagons alternately, corresponding to single and double bonds.

The figure clearly shows that in the diamond structure each atom is surrounded by four neighbors in a tetrahedral arrangement. The coordination number is therefore 4. This implies that space filling is much less efficient than in previously discussed structures. When this structure is built up of touching spheres of radius r, the relation between the edge of the Bravais cell and the radius is $a\sqrt{3}/4 = 2r$. Since each Bravais cell contains eight atoms, the packing fraction is

$$8\,\frac{4}{3}r^3\pi\left(\frac{8r}{\sqrt{3}}\right)^{-3} = \frac{\sqrt{3}}{16}\pi = 0.340\,. \tag{7.2.15}$$

This structure is realized in a particular form of carbon, diamond, as well as silicon, germanium, and gray tin (α-Sn). These elements are all in group 14 (IVA) of the periodic table where – as discussed in Chapter 4 – tetrahedral coordination is due to the covalent bonds created by the sp^3 hybrid orbitals.

This structure can also be considered to consist of two interpenetrating face-centered cubic Bravais lattices, displaced by $\frac{1}{4}a(\hat{\boldsymbol{x}}+\hat{\boldsymbol{y}}+\hat{\boldsymbol{z}})$ relative to one another. The obtained *diamond lattice* is not a Bravais lattice. It should be emphasized: even though the crystal is made up of a single kind of atom, it is impossible to choose a primitive cell with a monatomic basis that can serve to generate the entire crystal.

Concerning translational symmetries, the diamond lattice has the same primitive vectors as a face-centered cubic lattice, thus its reciprocal lattice is the same as that of an fcc lattice. The Brillouin zones are also identical: truncated octahedra.

On the other hand, rotation and reflection symmetries of the cube are not fully preserved. Inversion symmetry is lost because of the atoms in the octants. Similarly, reflections σ_x, σ_y, σ_z, fourfold rotations C_{4x}, C_{4y}, C_{4z} and twofold rotations around the face diagonals are no longer symmetries, either. In each case, however, the crystal can be brought into coincidence with itself by an additional translation along the space diagonal through one quarter of its length (that is, by the vector $\frac{1}{4}\frac{1}{4}\frac{1}{4}$).[6] If instead of the origin we were to choose the point $\frac{1}{8}\frac{1}{8}\frac{1}{8}$ as the fixed point of the point-group operations, inversion would remain a symmetry but invariance under other rotations would be broken – unless followed by a translation through one quarter of the space diagonal. Thus there are screw axes and glide planes among the symmetry elements of the diamond structure. The space group is $Fd\bar{3}m$ (O_h^7).

When the two sublattices are made up of two different kinds of atom, as shown in Fig. 7.16(b), a sphalerite (B3) structure is obtained. In sphalerite – a polymorph of ZnS – the basis contains two ions, Zn^{2+} at 000 and S^{2-} at

[6] This translation is neither along the screw axis nor in the glide plane. However, the same name may be used for convenience, as the contradiction with the definition of these symmetry operations is only apparent. See footnote on page 159. E.g., it is readily seen that the diamond lattice has a "pure" fourfold screw axis parallel to the z-axis if it is chosen to go through the point $\frac{1}{4}\frac{1}{2}0$.

$\frac{1}{4}\frac{1}{4}\frac{1}{4}$. The coordinates for the eight ions in the Bravais cell are

$$\text{Zn}: \quad 000,\ 0\tfrac{1}{2}\tfrac{1}{2},\ \tfrac{1}{2}0\tfrac{1}{2},\ \tfrac{1}{2}\tfrac{1}{2}0\,, \qquad \text{S}: \quad \tfrac{1}{4}\tfrac{1}{4}\tfrac{1}{4},\ \tfrac{1}{4}\tfrac{3}{4}\tfrac{3}{4},\ \tfrac{3}{4}\tfrac{1}{4}\tfrac{3}{4},\ \tfrac{3}{4}\tfrac{3}{4}\tfrac{1}{4}\,.$$

Compared to the diamond structure, the symmetry is lower. Inversion is no longer a symmetry, even when combined with a translation. The space group is therefore $F\bar{4}3m$ (T_d^2). Apart from the prototype, ZnS, several other compounds of composition AB crystallize in this form, many of which are important in semiconductor technology. For example, A = Zn, Cd, Hg, and B = Se, Te, S, or A = Ga, Al, In, and B = P, As, Sb.

Starting with the sphalerite structure, the addition of a third kind of atom to each empty octahedral site of face-centered Bravais cell leads to $C1_b$ structure (in the Pearson notation: cF12). The prototype is AgAsMg. The space group is again $F\overline{4}3m$ (T_d^2). This arrangement is also called a *half-Heusler structure* as it can be conceptually derived from a Heusler alloy ($AlMnCu_2$) by leaving half of the tetrahedral sites empty.

Contrarily, the space group $Fd\bar{3}m$ (O_h^7) of the diamond structure can be preserved by a suitable choice of filling its empty octant sites. This occurs in the C15 or Cu_2Mg structure (Fig. 7.17), the prototype for a class of Laves phase alloys. Just like in diamond, the eight magnesium atoms are arranged tetrahedrally in the Bravais cell: at the vertices and face centers plus at the centers of four of eight octants, while the sixteen Cu atoms are grouped four by four, in tetrahedra around the centers of the other four octants. The edge length of these small tetrahedra is such that the four Cu atoms closest to the center of the Bravais cell form an identical tetrahedron.

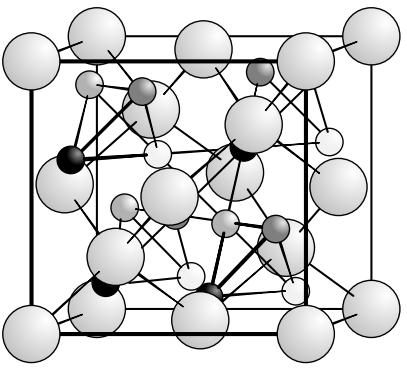


Fig. 7.17. Cu_2Mg, the prototype of C15 (cF24) structures

The same space group $Fd\bar{3}m$ (O_h^7) is present in the spinel (or $H1_1$) structure, named after the mineral $MgAl_2O_4$. The general composition of the compounds that crystallize in this structure is $X^{2+}Y_2^{3+}O_4$, where X stands for Mg, Fe, Zn, Mn, etc., and Y for Al, Fe, or Cr. Accordingly, one speaks of aluminate, ferrite, and chromite spinels. This structure type is of particular

importance because antiferromagnetic and ferrimagnetic ferrites that are extensively used in electrotechnology crystallize in this geometry. Oxygen atoms are located at the vertices and face centers of the cubic lattice, X^{2+} cations occupy one-eighth of the tetrahedral interstices, and Y^{3+} cations – one-half of the octahedral ones. As the cations are not at the same positions in each cube, the Bravais cell contains eight cubes and eight molecules – 56 atoms in all (cF56). In the inverse spinel structure divalent cations are at octahedral sites and half of the trivalent cations are at tetrahedral sites. Such an inverse spinel structure is observed in magnetite, in which X^{2+} and Y^{3+} are di- and trivalent ions of the same element, iron.

7.3 Hexagonal Crystal Structures

We have already seen (Fig. 7.12) that atoms are arranged in a hexagonal array in the (111) plane of a face-centered cubic crystal. Within the plane, each atom is surrounded by six neighbors, which is the closest packing in two dimensions. When such a hexagonal planar lattice is repeated at regular intervals along the perpendicular direction, as shown in 7.18(a), a *simple hexagonal lattice* is obtained.

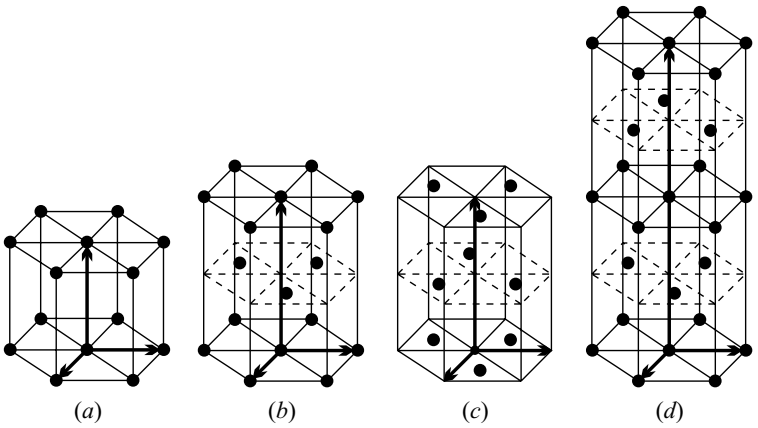


Fig. 7.18. Hexagonal structures: (a): simple; (b) and (c): close-packed; (d) double close-packed

Denoting the lattice constant in the hexagonal plane by a and in the perpendicular direction by c, a possible choice for the primitive vectors in Cartesian coordinates is

$$\boldsymbol{a}_1 = \frac{\sqrt{3}}{2} a\hat{\boldsymbol{x}} - \frac{1}{2} a\hat{\boldsymbol{y}}\,, \qquad \boldsymbol{a}_2 = a\hat{\boldsymbol{y}}\,, \qquad \boldsymbol{a}_3 = c\hat{\boldsymbol{z}}\,. \tag{7.3.1}$$

Matrices A and B are

$$A_{\text{hex}} = \begin{pmatrix} \sqrt{3}a/2 & 0 & 0 \\ -a/2 & a & 0 \\ 0 & 0 & c \end{pmatrix}, \qquad B_{\text{hex}} = 2\pi \begin{pmatrix} 2/(\sqrt{3}a) & 0 & 0 \\ 1/(\sqrt{3}a) & 1/a & 0 \\ 0 & 0 & 1/c \end{pmatrix}. \tag{7.3.2}$$

The primitive vectors of the reciprocal lattice are then

$$\boldsymbol{b}_1 = \frac{4\pi}{\sqrt{3}a}\hat{\boldsymbol{x}}\,, \quad \boldsymbol{b}_2 = \frac{2\pi}{\sqrt{3}a}\hat{\boldsymbol{x}} + \frac{2\pi}{a}\hat{\boldsymbol{y}}\,, \quad \boldsymbol{b}_3 = \frac{2\pi}{c}\hat{\boldsymbol{z}}\,. \tag{7.3.3}$$

As shown in Fig. 7.19, these vectors also generate a hexagonal lattice with lattice parameters $a^* = 4\pi/\sqrt{3}a$ and $c^* = 2\pi/c$, however, the reciprocal-lattice primitive vectors are rotated around the z-axis with respect to their direct-lattice counterparts, and their angle is 60° instead of 120°.

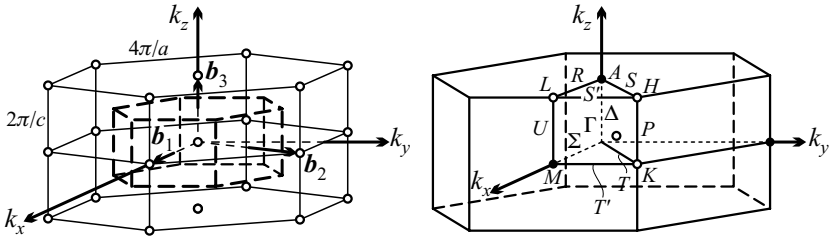


Fig. 7.19. The reciprocal lattice of a hexagonal lattice and its Brillouin zone with special points

The lattice points of a simple hexagonal lattice are directly above each other along the z direction. When this lattice is decorated with a single atom of radius r, the lattice is most densely filled when the atoms are in contact with each other both in the hexagonal plane and in the perpendicular direction. This occurs for $a = c = 2r$, and thus the volume of the hexagonal cell is

$$\frac{\sqrt{3}}{2}a^2c = 4\sqrt{3}r^3\,. \tag{7.3.4}$$

The packing fraction is therefore

$$\frac{4}{3}r^3\pi\left(4\sqrt{3}r^3\right)^{-1} = \frac{\pi}{3\sqrt{3}} = 0.605\,. \tag{7.3.5}$$

This is 20% smaller than the atomic packing factor of the fcc structure. This is why the above simple hexagonal structure may not occur naturally.

Space filling is more efficient if the planes of hexagonally arranged atoms are not stacked directly on top of each other, but the atomic positions are slightly shifted in every second layer, and only the next-nearest layers come

exactly above each other, as shown in Fig. 7.18(b). It is readily established that in the closest-packed arrangement of two subsequent layers each atom in the upper layer sits exactly into the dip among three touching spheres of the lower layer. Then the height difference between the centers of the spheres of the two layers is $\sqrt{8/3}r$. Since the lattice parameter in the hexagonal plane is $2r$ and the periodicity in the perpendicular direction is twice the repeat distance of the layers, close packing is achieved if $c/a = \sqrt{8/3} \approx 1.633$. As the cell now contains two atoms, the packing fraction is

$$2\sqrt{\frac{3}{8}}\,\frac{\pi}{3\sqrt{3}} = \frac{\pi}{3\sqrt{2}} = 0.740\,. \tag{7.3.6}$$

This is the same as the atomic packing factor of the close-packed fcc lattice. The structure with this particular c/a value is called the *hexagonal close-packed (hcp) structure*. It is also called Mg or A3 structure.

This structure can also be viewed as the superposition of two simple hexagonal Bravais lattices, with the sublattices displaced by

$$\frac{2}{3}\boldsymbol{a}_1 + \frac{1}{3}\boldsymbol{a}_2 + \frac{1}{2}\boldsymbol{a}_3 = \frac{\sqrt{3}}{3}a\hat{\boldsymbol{x}} + \frac{1}{2}c\hat{\boldsymbol{z}} \tag{7.3.7}$$

or

$$\frac{1}{3}\boldsymbol{a}_1 + \frac{2}{3}\boldsymbol{a}_2 + \frac{1}{2}\boldsymbol{a}_3 = \frac{\sqrt{3}}{6}a\hat{\boldsymbol{x}} + \frac{1}{2}a\hat{\boldsymbol{y}} + \frac{1}{2}c\hat{\boldsymbol{z}} \tag{7.3.8}$$

with respect to one another. Such a crystal has a diatomic basis and thus no longer possesses each rotational symmetry of the hexagonal lattice. To facilitate the visualization of symmetries, atomic positions are shown in a displaced coordinate system in Fig. 7.18(c). In an oblique coordinate system particularly well suited to hexagonal symmetry, with the x- and y-axes making an angle of 120° with each other, the coordinates of the two atoms of the primitive cell are $\frac{2}{3}\,\frac{1}{3}0$ and $\frac{1}{3}\,\frac{2}{3}\,\frac{1}{2}$. It is readily seen that the z-axis is not a sixfold rotation axis now but a sixfold screw axis: if a rotation through 60° around the z-axis is followed by a translation along the z-axis through $c/2$, the crystal is taken into itself. The space group of a close-packed hexagonal crystal is therefore $P6_3/mmc$ (D^4_{6h}).

A somewhat more complicated structure with the same symmetry is obtained by another way of stacking hexagonal layers. Four simple hexagonal sublattices may be superposed in such a way that the first, second, and third are displaced by $\frac{2}{3}\boldsymbol{a}_1 + \frac{1}{3}\boldsymbol{a}_2 + \frac{1}{4}\boldsymbol{a}_3$, $\frac{1}{2}\boldsymbol{a}_3$, and $\frac{1}{3}\boldsymbol{a}_1 + \frac{2}{3}\boldsymbol{a}_2 + \frac{3}{4}\boldsymbol{a}_3$ relative to the first. This leads to the *double hexagonal close-packed (dhcp) structure*, also called La structure or A3′ structure. Ideal close-packing occurs for $c/a = 2\sqrt{8/3} = 3.266$.

Over 30 elements crystallize naturally in hcp or dhcp structure, e.g., Be, Cd, Co, He, La, Mg, Os, Sc, Ti, Y, Zn, Zr, as well as several lanthanoids and actinoids. In these materials the ratio of the two lattice constants is not the ideal value $c/a = 1.633$ or its double but it is usually very close to one of

them. At high pressures, helium also crystallizes in a hexagonal structure, and packing is almost ideal: $c/a = 1.631$. The lattice constant ratio is very close to the ideal value also in alkali metals and the prototype Mg ($c/a = 1.637$ for Li, 1.634 in Na, and 1.624 in Mg). Lanthanum crystallizes in a dhcp structure with $c/a = 2 \times 1.6125$. Much larger discrepancies occur in Be ($c/a = 1.568$), as well as in Zn and Cd (1.856 and 1.886). Nevertheless even these structures are called close-packed.

In these structures the close-packed layers are stacked on top of each other as close as possible. As shown in Fig. 7.20, there are two different ways to stack a hexagonal close-packed layer on another by exploiting the dips among the atoms. If the arrangement of the first plane is called A, that of the next is called B or C, depending on whether it is like part (a) or (b) in Fig. 7.20.

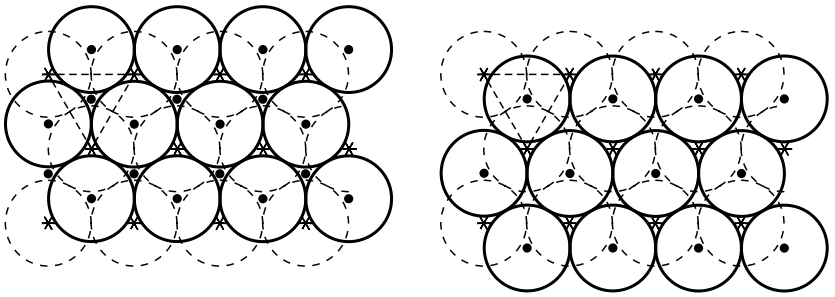


Fig. 7.20. Two possible ways of stacking hexagonal close-packed layers. Dashed-line circles (spheres) centered at asterisks ($*$) mark the atoms of the lower layer. Relative to this layer (A), two arrangements – shown in (a) and (b) – are possible for the centers ($\bullet$) of the spheres in the second layer. They are called type B and type C layers

The close-packed layers in a hcp crystal are stacked in such a way that two types alternate: $ABAB\ldots$ or $ACAC\ldots$. The stacking order in a double hexagonal close-packed structure is $ABACABAC\ldots$. Looking back at Fig. 7.12, the stacking order of the close-packed planes is seen to be $ABCABC\ldots$ along the [111] direction of the face-centered cubic crystal. The nearest neighbors of a selected atom are shown in Fig. 7.21 for both arrangements. In the fcc structure, the twelve neighbors are on the vertices of a cuboctahedron, while in the hcp structure they are at the vertices of an anticuboctahedron (also called triangular orthobicupola).

One may choose more complicated bases in the hexagonal structure, too. Figure 7.22 shows two such examples. Among two-component compounds of composition AB the NiAs structure or B8 structure is quite common. Atoms of one kind sit at positions 000 and $00\frac{1}{2}$, while atoms of the other at $\frac{1}{3}\frac{2}{3}\frac{1}{4}$ and $\frac{2}{3}\frac{1}{3}\frac{3}{4}$. This structure may be regarded as a modification of the double hexag-

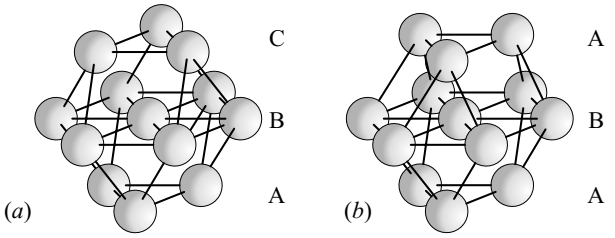


Fig. 7.21. The arrangement of the 12 nearest neighbors of an atom in a (*a*) face-centered cubic; (*b*) hexagonal close-packed structure

onal close-packed structure, where different kinds of atoms sit in alternate layers, in such a way that in spite of the unequal radii relatively close packing is maintained with a stacking order $ABAC\,ABAC\ldots$.

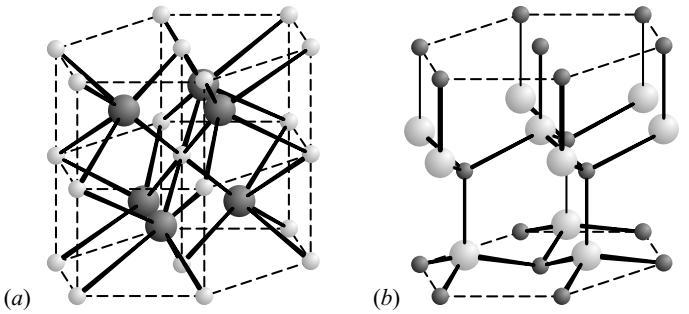


Fig. 7.22. Hexagonal crystal structures: (*a*) NiAs (B8) structure; (*b*) wurtzite (B4) structure

In addition to its cubic form, zinc sulfide (ZnS) may also crystallize in hexagonal geometry; it is then called *wurtzite*. In the wurtzite or B4 structure zinc atoms are located at $\frac{1}{3}\frac{2}{3}0$ and $\frac{2}{3}\frac{1}{3}\frac{1}{2}$, and sulfur atoms at $\frac{1}{3}\frac{2}{3}z$ and $\frac{2}{3}\frac{1}{3}(\frac{1}{2}+z)$, with $z \approx 3/8$. Zinc atoms are coordinated tetrahedrally by sulfur atoms, similarly to the cubic ZnS structure, sphalerite, shown in Fig. 7.16; however, the two forms differ in the relative orientation of tetrahedra in subsequent layers. Figure 7.22(*b*) shows this structure in a displaced coordinate system where a zinc atom is at the origin. Here, too, hexagonal layers of zinc and sulfur atoms alternate, however, the stacking order is now $AABB\,AABB\ldots$. Like above, the sixfold rotation axis is replaced by a sixfold screw axis and inversion is no longer a symmetry unless it is followed by a translation. The space group is therefore $P6_3mc$ (C_{6v}^4).

The hexagonal planes may be stacked in various other ways, leading to even longer periodicities in the structure. Structures with a periodicity of five ($ABCAB\,ABCAB\ldots$) or six ($ABCACB\,ABCACB\ldots$ or $ABABAC$

ABABAC ...) layers are equally possible. In a particular form of SiC a unit that extends over almost 600 layers is repeated.

The structure of graphite merits separate discussion (Section 7.5). Here we just mention that it can be regarded as four interpenetrating hexagonal Bravais lattices with origins at 000, $\frac{1}{3}\frac{2}{3}0$, $00\frac{1}{2}$, and $\frac{2}{3}\frac{1}{3}\frac{1}{2}$. This leads to the graphite structure shown in Fig. 7.23(a). In each plane carbon atoms form a honeycomb lattice, however, because of the displacement of subsequent layers only atoms in every second layer are exactly above each other.

7.4 Typical Sizes of Primitive Cells

We have not yet discussed the sizes of primitive cells, i.e., the characteristic distance of periodicity in the crystal. To fill this gap, we have listed the Bravais cell dimensions and nearest-neighbor distances for some of the cubic, tetragonal, and hexagonal structures that we had encountered. In line with common practice, lattice constants are given in angstroms – although using nm and pm units for atomic dimensions is more and more common. $1\,\text{Å} = 0.1\,\text{nm} = 100\,\text{pm}$.

As it can be seen in the table, atomic distances are a few angstroms in these simple ordered crystalline structures. As we shall see at the end of this chapter, these distances are often in good agreement with the values determined from atomic or ionic radii assuming close packing.

In more complicated structures the lattice constant may be much larger. The lattice constant of the fcc Bravais cell of fullerite is 14.16 Å at room temperature. The distance between centers of neighboring fullerene (C_{60}) molecules of diameter 7.10 Å is 10.02 Å; within a molecule the distance between two adjacent carbon atoms is 1.46 Å for single and 1.40 Å for double bonds. This is hardly different from the C–C distances in the hexagonal rings of graphite.

The order of magnitude of these distances is of great importance since their experimental determination requires methods that work well on such scales. These methods will be presented in the next chapter.

7.5 Layered and Chain-Like Structures

Most of the examples that we have encountered so far are cubic or hexagonal crystals. These types occur most frequently in nature, even though over a hundred other categories are listed in crystallographic tables. Here we shall consider a few special structures that are of particular interest for physicists because of their properties directly related to structure.

In cubic structures atoms are spaced at equal distances in the three spatial directions. In hexagonal and tetragonal structures one direction has a privileged status over the two others, however, atomic spacing along the sixfold

Table 7.2. Size of the Bravais cell and nearest-neighbor distance in some common cubic, tetragonal, and hexagonal crystal structures

Name of crystal	Structure type	Size of the Bravais cell		Nearest-neighbor distance
Po	A_h	$a = 3.366$ Å		3.366 Å
Ag	A1	$a = 4.086$ Å		2.889 Å
Al	A1	$a = 4.049$ Å		2.863 Å
Au	A1	$a = 4.078$ Å		2.884 Å
Ca	A1	$a = 5.588$ Å		3.951 Å
Cu	A1	$a = 3.615$ Å		2.556 Å
Kr	A1	$a = 5.646$ Å		3.992 Å
Ne	A1	$a = 4.462$ Å		3.155 Å
Ni	A1	$a = 3.523$ Å		2.491 Å
Pd	A1	$a = 3.890$ Å		2.751 Å
Pt	A1	$a = 3.924$ Å		2.775 Å
Cr	A2	$a = 2.884$ Å		2.498 Å
Cs	A2	$a = 6.067$ Å		5.254 Å
Fe	A2	$a = 2.867$ Å		2.483 Å
K	A2	$a = 5.321$ Å		4.608 Å
Na	A2	$a = 4.210$ Å		3.646 Å
W	A2	$a = 3.165$ Å		2.741 Å
Mg	A3	$a = 3.209$ Å	$c = 5.210$ Å	3.197 Å
Na	A3	$a = 3.767$ Å	$c = 6.154$ Å	3.768 Å
Zn	A3	$a = 2.664$ Å	$c = 4.949$ Å	2.665 Å
La	A3′	$a = 3.770$ Å	$c = 12.159$ Å	3.739 Å
C	A4	$a = 3.567$ Å		1.545 Å
Si	A4	$a = 5.431$ Å		2.352 Å
Ge	A4	$a = 5.657$ Å		2.450 Å
C	A9	$a = 2.461$ Å	$c = 6.709$ Å	1.426 Å
NaCl	B1	$a = 5.640$ Å		2.820 Å
CsCl	B2	$a = 4.113$ Å		2.908 Å
ZnS	B3	$a = 5.411$ Å		2.343 Å
ZnS	B4	$a = 3.822$ Å	$c = 6.261$ Å	2.343 Å
NiAs	$B8_1$	$a = 3.619$ Å	$c = 5.034$ Å	2.439 Å
CaF_2	C1	$a = 5.462$ Å		2.365 Å
Cu_2Mg	C15	$a = 7.048$ Å		2.492 Å
AuCu	$L1_0$	$a = 3\,966$ Å	$c = 3\,673$ Å	2.703 Å
Cu_3Au	$L1_2$	$a = 3.748$ Å		2.651 Å
Mn_3Pt	$L1_2$	$a = 3.833$ Å		2.710 Å

or fourfold rotation axis and in the perpendicular directions are of the same order. In some materials much larger discrepancies may arise between inequivalent directions. This may be due to the shape of the molecules making up the crystal – or to the nature of the bonds that determine the structure.

Besides diamond, the lattice constants of another allotrope of carbon, graphite are also listed in Table 7.2. As it can be seen from the illustration of the structure, Fig. 7.23(a), in the plane perpendicular to the sixfold axis carbon atoms form a honeycomb lattice in which each atom is surrounded by three neighbors at a distance of 1.42 Å. The separation between adjacent layers is more than its double, 3.35 Å.

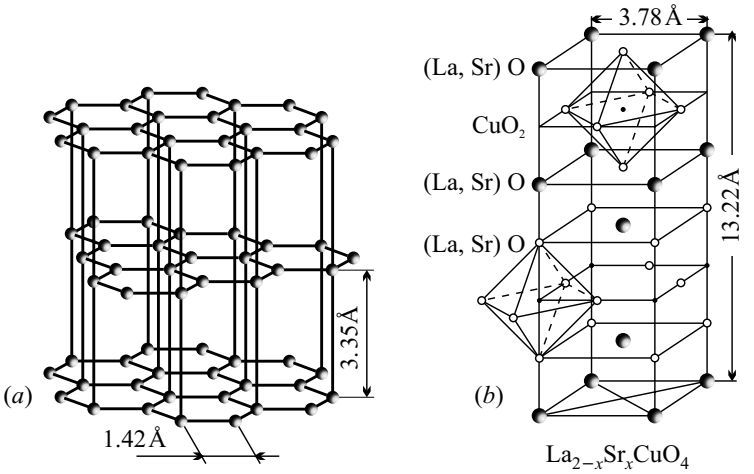
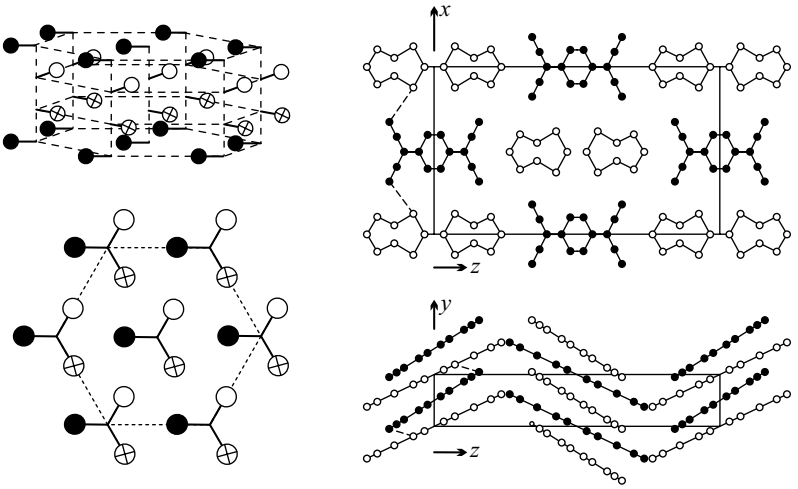


Fig. 7.23. The layered structure of (a) graphite; (b) $La_{2-x}Sr_xCuO_4$

Solid-state physics is more and more concerned with the study of such structures, in which the separation of atoms within a plane is much smaller than the separation between the layers. Figure 7.23(b) shows the primitive cell of another material with layered structure, $La_{2-x}Sr_xCuO_4$, the first example for high-T_c superconductors. Lattice parameters can be read off from the figure. In this structure, CuO_2 units form layers; the other atoms are located between these layers.

Another fairly common situation is that atoms are close together only along one direction, not in a plane. Figure 7.24 shows two structures in which the constituent atoms and molecules of the crystal form chains. The distances between the chains are quite large compared to those within the chains, and thus the bonds between the chains are weak. In an allotrope of selenium – γ-Se, the prototype of A8 structure – the crystal structure is hexagonal with lattice constants $a = 4.36$ Å and $c = 4.95$ Å. In the oblique coordinate system suited to hexagonal geometry, Se atoms are at $x00$, $\bar{x}\bar{x}\frac{1}{3}$, and $0x\frac{2}{3}$, where

$x \approx 0.22$. The crystal is invariant under the combined operation of a rotation through 120° around the c-axis and a translation through one-third the lattice constant. The space group of the crystal is thus $P3_121$ (D_3^4). As it is shown in Fig. 7.24(a), the distance between neighboring atoms along the spiral is smaller than the separation of the spirals. The structure is thus similar to a set of twisted chains.



(a) (b)

Fig. 7.24. (a) The hexagonal A8 structure of γ-Se and its projection on the (001) plane. Atoms marked by •, ∗ and ∘ signs are on planes $c/3$ apart. (b) The projection of the atoms of HMTTF-TCNQ to the (010) and (100) planes

Chain-like structures arise when large flat molecules are stacked relatively tightly and the distance between the chains is larger than the separation of the molecules in them. Figure 7.24(b) shows two projections of the orthorhombic primitive cell of HMTTF-TCNQ – a material showing metallic characteristics in spite of being built up of two organic molecules, HMTTF and TCNQ. The space group is $Pmna$ (D_{2h}^7). The lattice constants of the primitive cell are $a = 12.470$ Å, $b = 3.906$ Å, and $c = 21.602$ Å. HMTTF cations occupy sites 000 and $\frac{1}{2}\frac{1}{2}\frac{1}{2}$, while TCNQ anions are found at $\frac{1}{2}\frac{1}{2}0$ and $00\frac{1}{2}$. As the figure shows, the molecules are not precisely in the ac-plane but are inclined at 23.8° and 34.2° with respect to it. Thus in the direction of the b-axis the separation of TCNQ molecules is 3.23 Å, while that of HMTTF molecules is 3.57 Å. These distances are much smaller than intermolecular distances in the ac-plane, thus one is justified to say that HMTTF and TCNQ molecules form chains along the b-axis, and the interaction between the chains is relatively weak.

One possible consequence of the chain-like structure is the almost vanishing overlap of the electron clouds between the chain. In layered structures the

overlap is usually much weaker between layers than within them. This gives rise to highly anisotropic physical properties. For example, in crystals built up of HMTTF and TCNQ (or similar pairs of) molecules electric conductivity is much larger in one direction than in the other two. $La_{2-x}Sr_xCuO_4$ and other crystals with CuO_2 planes are much better conductors in the plane than in the perpendicular direction. This property may in itself be crucial for some applications. It is also of great importance at the phenomenological level: as we shall see, fundamentally new phenomena may be observed in materials in which the system of electrons is practically one- or two-dimensional. This explains why the study of chain-like and layered structures has become one of the hottest research topics in solid-state physics.

7.6 Relationship Between Structure and Bonding

We saw in Section 4.4 that unlike other types of bonds, covalent bonds are highly directional. This has a strong influence on the structures occurring for each particular type of chemical bonding.

7.6.1 The Structure of Covalently Bonded Solids

As mentioned in Subsection 4.4.7 and shown in Fig. 4.11, sp^3 hybrid orbitals can create bonds in the directions of the four vertices of a tetrahedron. This is the underlying reason for the tetrahedral coordination of atoms in the diamond structure. Tetrahedral arrangement of nearest neighbors is often observed in two-component covalent compounds, too. The commonest are the sphalerite and wurtzite structures. In such tetrahedrally bonded covalent crystals a covalent radius can be determined for each atom from the requirement that the sum of covalent radii should give the *bond length* (also called *bond distance*) – i.e., the distance between the atoms in the covalent bond. The covalent radius derived from the bond length may also be introduced for covalently bonded materials with other structures, however, the lengths are different for atoms linked by single and double covalent bonds. Covalent radii – with the above uncertainty – are listed in Table 7.3 for some elements.

If besides the s-electron only two p-electrons (p_x and p_y) participate in bonding, then the following sp^2 hybrid wavefunctions arise:

$$\begin{aligned}
\psi_1 &= \frac{1}{\sqrt{3}}\phi_{2s} + \sqrt{\frac{2}{3}}\phi_{2p_x}\,, \\
\psi_2 &= \frac{1}{\sqrt{3}}\phi_{2s} - \frac{1}{\sqrt{6}}\phi_{2p_x} + \frac{1}{\sqrt{2}}\phi_{2p_y}\,, \\
\psi_3 &= \frac{1}{\sqrt{3}}\phi_{2s} - \frac{1}{\sqrt{6}}\phi_{2p_x} - \frac{1}{\sqrt{2}}\phi_{2p_y}.
\end{aligned} \tag{7.6.1}$$

As a consequence of the form of s and p wavefunctions shown in Fig. 4.10, these hybrid states give high electron densities in three directions of the (x, y)

Table 7.3. Covalent radii in Å for elements that participate in covalent bonds

	B	C	N	O
	0.81	0.77	0.70	0.66
	Al	Si	P	S
	1.25	1.17	1.10	1.04
Zn	Ga	Ge	As	Se
1.25	1.25	1.22	1.21	1.17
Cd	In	Sn	Sb	Te
1.41	1.50	1.40	1.41	1.37
Hg	Tl	Pb	Bi	Po
1.44	1.55	1.54	1.46	1.46

plane, at 120° degrees to each other. This is illustrated in Fig. 7.25(a). These hybrid bonding orbitals give rise to the two-dimensional honeycomb-like network shown in Fig. 7.25(b).

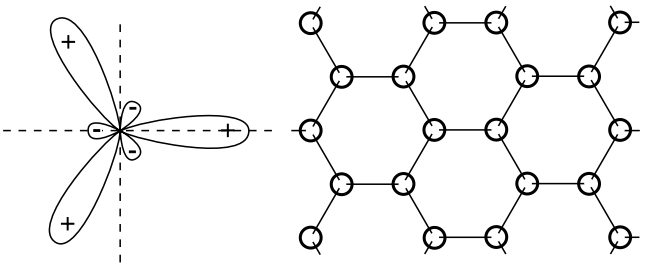


Fig. 7.25. (a) Spatial distribution of electrons in states characterized by sp^2 hybrid wavefunctions. (b) Two-dimensional network arising from these bonds

This is the case for an allotrope of carbon, graphite, where only two 2p-electrons hybridize with one 2s-electron. In addition to the three electrons that participate in saturated covalent bonds, a fourth electron makes a nonlocalized and unsaturated π bond with the three neighbors. This electron is responsible for the finite conductivity of graphite. The spatial structure of graphite – the prototype of A9 structure – is shown in Fig. 7.23(a): covalently bonded hexagonal planes are held together by weak van der Waals forces. The same planar network is found in trivalent As.

Besides s- and p-states, one or more 3d-states may also participate in the formation of hybrid bonding orbitals. Considering the form of d-states, illustrated in Fig. 6.1, it can be demonstrated that in a dsp^2 hybrid the wavefunctions

$$\begin{aligned}\psi_{1,2} &= \frac{1}{2}\left[\phi_s \pm \sqrt{2}\phi_{p_x} + \phi_{d_{x^2-y^2}}\right], \\ \psi_{3,4} &= \frac{1}{2}\left[\phi_s \pm \sqrt{2}\phi_{p_y} - \phi_{d_{x^2-y^2}}\right]\end{aligned} \tag{7.6.2}$$

can create bonds along the $\pm x$ and $\pm y$ directions of the (x, y) plane, in a tetragonal geometry.

In the d^2sp^3 hybrid ϕ_s, ϕ_{p_x}, ϕ_{p_y}, and ϕ_{p_z} states are combined with $\phi_{d_{z^2}}$ and $\phi_{d_{x^2-y^2}}$ to give

$$\begin{aligned}\psi_{1,2} &= \frac{1}{\sqrt{6}}\phi_s \pm \frac{1}{\sqrt{2}}\phi_{p_z} + \frac{1}{\sqrt{3}}\phi_{d_{z^2}}, \\ \psi_{3,4} &= \frac{1}{\sqrt{6}}\phi_s \pm \frac{1}{\sqrt{2}}\phi_{p_x} + \frac{1}{2}\phi_{d_{x^2-y^2}} - \frac{1}{\sqrt{12}}\phi_{d_{z^2}}, \\ \psi_{5,6} &= \frac{1}{\sqrt{6}}\phi_s \pm \frac{1}{\sqrt{2}}\phi_{p_y} - \frac{1}{2}\phi_{d_{x^2-y^2}} - \frac{1}{\sqrt{12}}\phi_{d_{z^2}}.\end{aligned} \tag{7.6.3}$$

In states associated with these functions electrons create bonds of octahedral configuration.

In covalently bonded solids each atom (molecule) is usually surrounded by the same number of nearest neighbors as the covalent bonds it can form. This is why the hexagonal and tetrahedral structures illustrated in Figs. 7.25 and 7.16 occur frequently. When an atom can form only two covalent bonds, a chain-like structure arises. An example for this, γ-selenium is shown in Fig. 7.24(a). The two covalent bonds do not lie along a straight line, which is why atoms are located along a spiral.

The strong directionality of bonds is preserved even when the covalently bonded elements make up a disordered, amorphous structure instead of a regular crystal. The bond lengths and angles are close to their values in a crystalline material, therefore these amorphous systems exhibit short-range order on atomic scales. On larger scales the order may disappear, as shown in Fig. 10.1.

7.6.2 Structures with Nondirectional Bonds

Unlike covalently bonded solids, the constituents of molecular crystals, ionic crystals, and metals with nondirectional van der Waals, ionic, or metallic bonds tend to arrange themselves as closely packed as possible. In the foregoing we have seen that the most effective filling of space is offered by face-centered cubic and hexagonal close-packed lattices; this is why most metals crystallize in one of these structures. Similar close-packed arrangements are found in molecular crystals, too.

The situation is similar for ionic crystals, although the different size of cations and anions plays an important role there. Although electron density decreases continuously with the distance from the ion core, considering the ions as rigid spheres of finite ionic radius r_i seems to give a fair approximation.

Because of the tendency to reach the energy minimum, the spheres will try to fill space in the most closely packed arrangement. One would expect that the nearest-neighbor distance is the sum of the ionic radii of an anion and a cation. As we shall see, for geometrical reasons this cannot be the case for arbitrary ionic radii.

Ionic radii cannot be determined unambiguously. Relying on the assumption that atoms are close-packed, several attempts have been made to estimate ionic radii from the lattice constant of the ionic crystal and the separation d between neighboring cations and anions using

$$d = r_{\mathrm{c}} + r_{\mathrm{a}}\,, \tag{7.6.4}$$

where r_{c} is the radius of the cation and r_{a} is the radius of the anion. V. M. Goldschmidt (1926) chose the radii of O^{2-} and F^{-} ions as reference, assuming the values: $r_{\mathrm{i}}(O^{2-}) = 1.32\,\text{Å}$ and $r_{\mathrm{i}}(F^{-}) = 1.33\,\text{Å}$. One year later (1927) L. C. Pauling worked out a system in which the values $r_{\mathrm{i}}(O^{2-}) = 1.40\,\text{Å}$ and $r_{\mathrm{i}}(F^{-}) = 1.36\,\text{Å}$ were chosen. To illustrate the variations of ionic radii in the periods and groups of the periodic table, the Pauling ionic radii of some singly and doubly charged ions are listed in Table 7.4.

Table 7.4. Pauling ionic radii (in Å) of some singly and doubly charged ions

Li^{+}	Be^{2+}			O^{2-}	F^{-}
0.60	0.31			1.40	1.36
Na^{+}	Mg^{2+}			S^{2-}	Cl^{-}
0.95	0.65			1.84	1.81
K^{+}	Ca^{2+}	Cu^{+}	Zn^{2+}	Se^{2-}	Br^{-}
1.33	0.99	0.96	0.74	1.98	1.95
Rb^{+}	Sr^{2+}	Ag^{+}	Cd^{2+}	Te^{2-}	I^{-}
1.48	1.13	1.26	0.97	2.21	2.16
Cs^{+}	Ba^{2+}	Au^{+}	Hg^{2+}		
1.69	1.35	1.37	1.10		

In reality, ionic radii may depend on the number of nearest neighbors (the coordination number). Data listed in Table 7.5 clearly show that when the radius of the same ion is determined from the lattice constants of different types of crystals, highly disparate values are obtained.

According to Table 4.4, cesium chloride has the largest Madelung constant among simple structures with a diatomic basis. Therefore when ions are considered as classical objects with a spherical charge distribution whose energy can be determined in terms of the energies of point charges, this structure should be more stable than sodium chloride, sphalerite, or wurtzite structures. Nevertheless the latter occur naturally quite frequently. To understand this, finite ionic radii are taken into account through the assumption that in

Table 7.5. Ionic radius (in Å) in environments with different coordination numbers, estimated from lattice constants

Coordination number	Al^{3+}	Mn^{2+}	Fe^{2+}	Fe^{3+}	Cu^{+}	Zn^{2+}	O^{2-}
4	0.39	0.66	0.63	0.49	0.60	0.60	
6	0.54	0.83	0.61	0.55	0.77	0.74	1.40
8		0.96	0.92	0.78		0.90	1.42

the energetically most favorable situation cations and anions are as close together as possible, i.e., in contact with each other inside the crystal. This is illustrated in Fig. 7.26, where the arrangement of touching ions is shown in the (110) plane of a CsCl structure for various values of the ionic radius ratio.

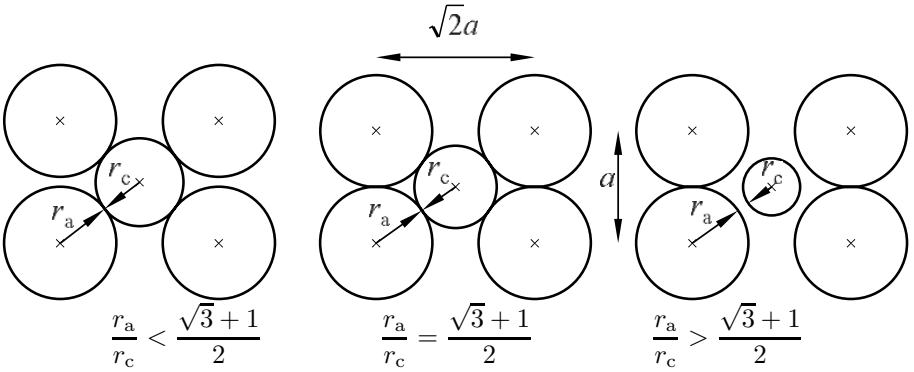


Fig. 7.26. Touching rigid spherical ions in the (110) plane of the cesium chloride structure, for various values of the anion–cation radius ratio

Taking the radius of the anions fixed and gradually reducing the radius of cations, anions and cations will be in contact along the space diagonal provided $2(r_a + r_c) = \sqrt{3}a$ and $2r_a \leq a$ both hold, that is, as long as

$$\frac{r_a}{r_c} \leq \frac{\sqrt{3}+1}{2} = 1.366\,. \tag{7.6.5}$$

Using the Pauling ionic radii, the ratio r_a/r_c for CsCl is 1.07, the above condition is thus met.

However, for Na^+ and Cl^- ions the same ratio is found to be 1.91. In a CsCl structure these anions and cations could not touch one another. In such cases, NaCl structure is energetically more favorable: although its Madelung constant is slightly smaller, it allows the two ions to be in contact. As illustrated in Fig. 7.27, cations and anions are in direct contact – and so sodium

chloride structure is stable – as long as $2(r_a + r_c) = a$ and $4r_a \leq \sqrt{2}a$, that is

$$\frac{r_a}{r_c} \leq \sqrt{2} + 1 = 2.415\,. \tag{7.6.6}$$

When the radius ratio exceeds this value, tetrahedrally coordinated sphalerite-type structure may occur. Similar considerations lead to the stability condition

$$\frac{r_a}{r_c} \leq 2 + \sqrt{6} = 4.444 \tag{7.6.7}$$

in this case.

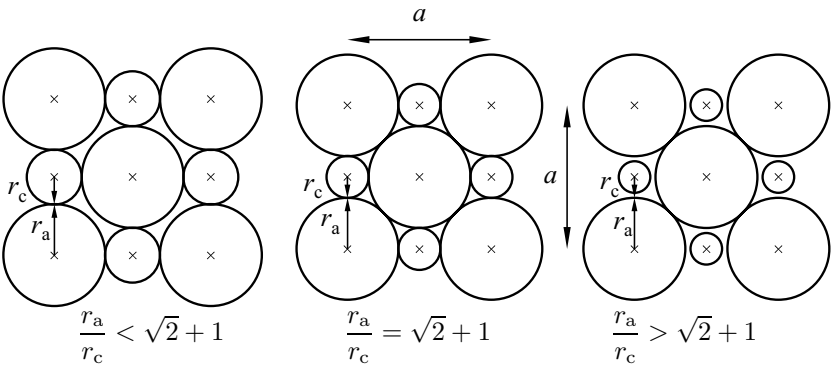


Fig. 7.27. Touching rigid spherical ions in the (110) plane of the sodium chloride structure, for various values of the anion–cation radius ratio

In the stable structure cations are visibly surrounded by eight anions when $1 < r_a/r_c < 1.366$, by six anions when $1.366 < r_a/r_c < 2.415$, and by just four anions when $2.415 < r_a/r_c < 4.444$. When the ratio of the ionic radii exceeds this value a trigonally coordinated planar configuration becomes stable. For even higher values of this ratio, $r_a/r_c > 2\sqrt{3} + 3 = 6.464$ linear coordination becomes stable.

Compared to the 1920s, we now have a much more profound understanding of the spatial distribution of electrons in ionic crystals. High-resolution diffraction measurements permit us to determine the density distribution of electrons. This is shown in Fig. 7.28 for rock salt.

The electron density reconstructed from measurements varies continuously, and it is not really spherically symmetric either – nevertheless it is reasonable to choose the minimum along the axis joining the two ions as the border between them. The "physical" ionic radius r_i^{ph} determined this way is listed in Table 7.6 for various ions at sites with coordination number six. "Physical" and Pauling ionic radii are seen to differ by as much as 0.1–0.2 Å, and should therefore be considered as just indications of the size of ions.

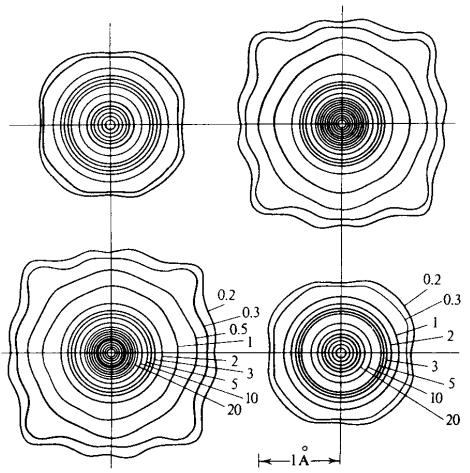


Fig. 7.28. Electron density in the (100) plane passing through atomic centers in a NaCl crystal, determined via the inverse Fourier transform of diffraction patterns [R. Brill, *Solid State Physics* **20**, 1 (1967)]

Table 7.6. "Physical" ionic radii (in Å) of some singly and doubly charged ions

Li^+	Be^{2+}			O^{2-}	F^-
0.94	0.59			1.26	1.16
Na^+	Mg^{2+}			S^{2-}	Cl^-
1.17	0.86			1.70	1.64
K^+	Ca^{2+}	Cu^+	Zn^{2+}	Se^{2-}	Br^-
1.49	1.14	0.91	0.88	1.84	1.80
Rb^+	Sr^{2+}	Ag^+	Cd^{2+}	Te^{2-}	I^-
1.63	1.32	1.29	1.09	2.07	2.05
Cs^+	Ba^{2+}	Au^+	Hg^{2+}		
1.86	1.49	1.51	1.16		

Further Reading

1. J. L. C. Daams, P. Villars, and J. H. N. van Vucht, *Atlas of Crystal Structure Types for Intermetallic Phases*, ASM International, Materials Park, OH (1991).
2. *Materials Science and Technology, A Comprehensive Treatment*, Edited by R. W. Cahn, P. Haasen, E. J. Kramer, Vol. 1. *Structure of Solids*, Volume Editor: V. Gerold, VCH Publishers Inc., New York (1993).
3. W. B. Pearson, *The Crystal Chemistry and Physics of Metals and Alloys*, Wiley-Interscience, New York (1972).

4. B. K. Vainshtein, V. M. Fridkin, and V. L. Indenbom, *Modern Crystallography*, Volume 2; *Structure of Crystals*, Third edition, Springer-Verlag, Berlin (2000).
5. P. Villars and L. D. Calvert, *Pearson's Handbook of Crystallographic Data for Intermetallic Phases*, Second edition, ASM International, Materials Park, Ohio (1991).
6. R. W. G. Wyckoff, *Crystal Structures*, Second Edition, John Wiley & Sons Inc., New York (1963).

8

Methods of Structure Determination

Symmetries of the regular crystalline arrangement of atoms manifest themselves in direction-dependent quantities such as elastic constants or electric conductivity. The measurement of such quantities therefore provides information about the symmetries of the crystal, and so, indirectly, about its structure. A full determination of the structure – the Bravais lattice as well as the atomic positions in the basis – nevertheless calls for a method by which the interior of the sample can be seen. This requires a radiation that penetrates relatively deeply into the material so that it should feel not only the atoms close to the surface. Its wavelength has to be comparable to interatomic distances, i.e., the dimensions of the primitive cell.

As we have seen, typical interatomic distances are a few angstroms, therefore the wavelength of the probing radiation has to be of the same order of magnitude. This corresponds to the X-ray region of the electromagnetic radiation spectrum. In 1912, using X-rays scattered by a crystal, M. von Laue[1] and two co-workers, W. Friedrich and P. Knipping obtained a diffraction pattern characteristic of interference. X-ray diffraction has since become the classic method of crystal-structure determination, due in a large part to the works of W. H. Bragg and W. L. Bragg.[2] The range of applications of X-ray diffraction has been extended substantially with the appearance of synchrotrons providing high-intensity radiation of high-energy γ photons.

Besides X-ray diffraction, elastic scattering of electrons or neutrons is also suitable for structure determination as it is fairly easy to produce sufficiently intense electron and neutron beams of appropriate wavelengths. By slowing down neutrons from fission reactors or spallation sources to thermal energies it is possible to produce neutrons in the wavelength range of a few angstroms. In the case of electrons, beams of the above wavelength are produced by accelerating them to energies on the order of 100 eV. In the section on experimental

[1] See footnote on page 2.
[2] See footnote on page 2.

methods we shall discuss the advantages and disadvantages of using each type of radiation.

Whether photons, electrons, or neutrons are scattered by condensed matter, besides its structure, other characteristics of the surface or bulk of the sample may also be revealed. This is because interaction with radiation may change the state of the sample, inducing a transition from its initial state into an(other) excited state. In this case the energy of the photons (electrons, neutrons) in the scattered beam is different from the energy of the particles in the incident beam. These inelastic processes provide information about the internal processes taking place inside the sample, and thus their observation opens the way to studying the excited states of the system. To this end, one has to measure the spatial and energy distribution of the radiation scattered by the sample. Structure determination is much simpler: it just requires the measurement of the spatial distribution of the scattered beam emerging from a much more intense process, elastic scattering (diffraction). In this chapter we first present the theory of diffraction, and then discuss the methods used to observe diffraction.

8.1 The Theory of Diffraction

Owing to the greater availability of X-ray sources, X-ray diffraction is more widely used for structure determination than neutron or electron diffraction. For this reason, in what follows we shall speak of X-ray diffraction, incident and scattered photons, although the condition presented below is valid for any type of radiation, it is concerned only with the wavelength. On the other hand, the penetration depth of the incident radiation into the sample and the relative intensity of the diffracted beam both depend on the character of the radiation. Thus the diffracted beam carries information either about the internal structure or the surface.

8.1.1 The Bragg and Laue Conditions of Diffraction

Shortly after the first X-ray diffraction patterns were recorded (1912), two different interpretations were proposed to account for them. W. L. Bragg (1913) advocated the view that atoms in crystals are arranged in parallel planes spaced at equal distances d, and X-rays are reflected specularly from them in accordance with the laws of reflection. As illustrated in Fig. 5.10, atoms can indeed be arranged into planes. Figure 8.1 shows a particular set of planes and the rays reflected from adjacent planes.

The phase difference between rays reflected from subsequent planes depends on the spacing between the planes. Constructive interference occurs when the phase difference between the rays reflected from adjacent planes is an integral multiple of 2π; otherwise the rays scattered by the planes interfere destructively. Thus scattered beams emerge only for some particular values of

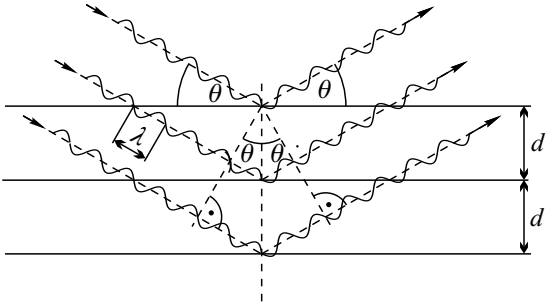


Fig. 8.1. Reflection of an X-ray beam of wavelength λ from a particular set of atomic planes separated by equal distances d. θ is the complement of the angle of incidence

the angle of incidence. On the other hand, the direction of the crystal planes can be chosen in infinitely many different ways, as shown in Fig. 5.10. Then the condition for constructive interference can be satisfied for several of them for a fixed incoming beam, and thus scattered beams can emerge in several directions.

To determine the condition for interference, assume that the incident beam makes an angle θ with the selected crystal plane.[3] According to the laws of reflection, the scattered beam makes the same angle with the plane, so the angle of deflection of the incident beam is 2θ. The path difference between two rays reflected from adjacent planes is $\Delta s = 2d \sin\theta$, as it is readily seen in the figure. The integral multiples of this distance appear for waves reflected from subsequent layers. In terms of the path difference Δs, constructive interference occurs when Δs is an integral multiple of the wavelength λ. Thus scattered beams emerge only when the condition

$$\boxed{2d \sin\theta = m\lambda} \tag{8.1.1}$$

is met by a family of crystal planes, where m is an integer. This is the *Bragg condition for diffraction*. The intensity of the reflected beam has sharp peaks in the corresponding directions. They are called *Bragg peaks*.

In LAUE's formulation the appearance of a sharp diffraction pattern is interpreted in terms of the interference of rays scattered by individual atoms rather than planes. Incident radiation is scattered by each atom, with an intensity that depends on the atomic species. The phase differences between scattered waves depend on atomic positions and the direction of the incident and scattered beams. To determine this phase difference assume that a radiation of wavelength λ is incident on the sample along the direction of the unit vector $\boldsymbol{n}$. Diffraction does not modify the wavelength of the ray, while its new direction will be denoted by $\boldsymbol{n}'$. Now consider two equivalent atoms in the

[3] In X-ray crystallography the above-defined θ is conventionally called the angle of incidence.

crystal. Obviously, when one is chosen as the origin, the position vector $\boldsymbol{R}_n$ of the other has to be a lattice vector of the crystal. The geometry of incident and scattered beams is illustrated in Fig. 8.2.

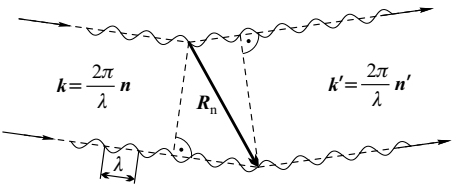


Fig. 8.2. Scattering of X-rays by two atoms of the crystal separated by a vector of the Bravais lattice

It can be immediately seen in the figure that the path difference between the waves scattered by the atoms at lattice point $\boldsymbol{R}_n$ and at the origin is

$$\Delta s = \boldsymbol{R}_n \cdot \boldsymbol{n} - \boldsymbol{R}_n \cdot \boldsymbol{n}' \,. \tag{8.1.2}$$

In terms of the wave vectors $\boldsymbol{k} = (2\pi/\lambda)\boldsymbol{n}$ and $\boldsymbol{k}'$, the corresponding phase difference is

$$\Delta\phi = \boldsymbol{R}_n \cdot \boldsymbol{k} - \boldsymbol{R}_n \cdot \boldsymbol{k}' \,. \tag{8.1.3}$$

The condition for constructive interference is that this phase difference be an integral multiple of 2π for any lattice vector $\boldsymbol{R}_n$ of the crystal, that is,

$$\boldsymbol{R}_n \cdot (\boldsymbol{k} - \boldsymbol{k}') = 2\pi m \,. \tag{8.1.4}$$

Recall (5.2.20): the product of any reciprocal-lattice vector $\boldsymbol{G}$ and any direct-lattice vector $\boldsymbol{R}_n$ is also an integral multiple of 2π. This means that the condition (8.1.4) for diffraction is equivalent to the requirement that $\boldsymbol{k}-\boldsymbol{k}'$ should be a reciprocal-lattice vector,

$$\boxed{\boldsymbol{k} - \boldsymbol{k}' = \boldsymbol{G} \,.} \tag{8.1.5}$$

This relationship is called the *Laue condition for diffraction*. For an incident beam of given direction and wavelength scattered waves emerge only in the directions that satisfy the above condition. We saw in Chapter 6 that for diffraction this condition is in fact the consequence of discrete translational symmetry, and as such independent of any properties of the scattered beam.

Equation (8.1.5) may be written in another useful form by rearranging the terms and squaring the sides:

$$|\boldsymbol{k}|^2 - 2\boldsymbol{k} \cdot \boldsymbol{G} + |\boldsymbol{G}|^2 = |\boldsymbol{k}'|^2 \,. \tag{8.1.6}$$

As scattering is elastic, $|\boldsymbol{k}| = |\boldsymbol{k}'|$, the condition for diffraction is

$$\boldsymbol{k} \cdot \boldsymbol{G} = \tfrac{1}{2}|\boldsymbol{G}|^2 , \tag{8.1.7}$$

which is equivalent to the requirement that $\boldsymbol{k} - \frac{1}{2}\boldsymbol{G}$ be perpendicular to $\boldsymbol{G}$. Thus the condition for diffraction is satisfied for incident beams whose wave vector points into a *Bragg plane*, i.e., the perpendicular bisecting plane of some vector $\boldsymbol{G}$ of the reciprocal lattice, as shown in Fig. 8.3.

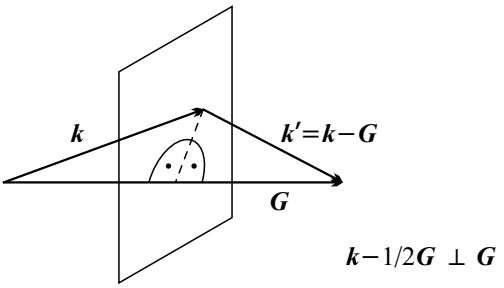


Fig. 8.3. Illustration of the Laue condition (8.1.7). The wave vector $\boldsymbol{k}$ of the incident beam has to point into the perpendicular bisecting plane of some vector $\boldsymbol{G}$ of the reciprocal lattice

Although derived from the Laue condition, this relation is often called the Bragg condition for diffraction. As it can be easily proved, the Bragg and Laue conditions are mathematically equivalent forms of the same requirement. Consider a situation in which the wave vectors of the incident and scattered beams satisfy condition (8.1.5) with a vector

$$\boldsymbol{G}_{hkl} = m(h\boldsymbol{b}_1 + k\boldsymbol{b}_2 + l\boldsymbol{b}_3) \tag{8.1.8}$$

of the reciprocal lattice, where h, k, and l are relatively prime. The geometry of vectors $\boldsymbol{k}$, $\boldsymbol{k}'$, and $\boldsymbol{G}_{hkl}$ is illustrated in Fig. 8.4.

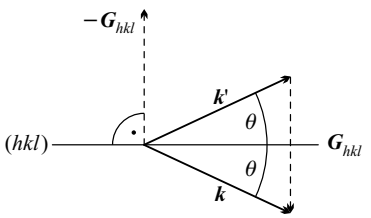


Fig. 8.4. Geometry of the wave vectors of the incident and scattered beams, the reciprocal-lattice vector satisfying the Laue condition, and the crystal plane perpendicular to it

As $\boldsymbol{k}$ and $\boldsymbol{k}'$ are of equal length, their angle bisector is perpendicular to $\boldsymbol{G}_{hkl} = \boldsymbol{k} - \boldsymbol{k}'$. Since the reciprocal-lattice vector $\boldsymbol{G}_{hkl}$ is perpendicular to the

lattice plane with Miller indices (hkl), the diffraction can be considered as if the incoming beam were incident on and then reflected from the (hkl) plane at an angle θ. From the triangle shown in the figure the Laue condition can also be written as

$$|\boldsymbol{G}_{hkl}| = 2|\boldsymbol{k}| \sin\theta \,. \tag{8.1.9}$$

According to (5.2.26), the length of the reciprocal-lattice vector $\boldsymbol{G}_{hkl}$ can be expressed in terms of the spacing d_{hkl} of the planes specified by the Miller indices (hkl). On account of the choice made in (8.1.8),

$$|\boldsymbol{G}_{hkl}| = m\frac{2\pi}{d_{hkl}} \,. \tag{8.1.10}$$

By eliminating the wave number $|\boldsymbol{k}|$ in favor of the wavelength, (8.1.9) takes the form

$$m\,\frac{2\pi}{d_{hkl}} = 2\frac{2\pi}{\lambda}\sin\theta \,. \tag{8.1.11}$$

This form of the Laue condition is obviously equivalent to the Bragg condition given in (8.1.1).

By measuring the angles where Bragg peaks appear, the spacings d_{hkl}, and from them the lattice parameters can be determined. For example, in cubic crystals the lattice constant of the Bravais cell is related to the spacing by

$$\frac{1}{d_{hkl}^2} = \frac{h^2 + k^2 + l^2}{a^2} \,, \tag{8.1.12}$$

while in orthorhombic crystals the three lattice constants can be determined from

$$\frac{1}{d_{hkl}^2} = \frac{h^2}{a^2} + \frac{k^2}{b^2} + \frac{l^2}{c^2} \,. \tag{8.1.13}$$

The unambiguous identification of the integers h, k, and l, and thereby the unambiguous determination of the lattice parameters calls for the measurement of the position of several interference peaks.

8.1.2 Structure Amplitude and Atomic Form Factor

In the previous subsection we derived the Bragg and Laue conditions from the wave nature of radiation using purely geometrical considerations, without any restrictions on the scattering potential. We only assumed that atoms are arranged in a periodic array, that the incident and scattered waves can be considered as plane waves, and that rays scattered by individual atoms or atomic planes interfere.

A more quantitative picture may be obtained by assuming that the interaction between the X-ray and the atom is given by a potential $U_{\mathrm{a}}(\boldsymbol{r})$. When the crystal has a monatomic basis, i.e., contains a single atom per primitive cell, the total potential is

$$U(\boldsymbol{r}) = \sum_{\boldsymbol{R}_n} U_{\mathrm{a}}(\boldsymbol{r} - \boldsymbol{R}_n)\,. \tag{8.1.14}$$

Since photons interact with the electron shell of the atom, the potential is proportional to the electron density,

$$U_{\mathrm{a}}(\boldsymbol{r} - \boldsymbol{R}_n) \propto \rho_{\mathrm{e}}(\boldsymbol{r} - \boldsymbol{R}_n)\,. \tag{8.1.15}$$

In the case of neutron scattering, neutrons interact with the nuclei or the magnetic moment of the electron cloud. Scattering by nuclei is described in terms of a local, Dirac delta-like Fermi pseudopotential,

$$U_{\mathrm{a}}(\boldsymbol{r} - \boldsymbol{R}_n) = -\frac{2\pi\hbar^2}{m_{\mathrm{n}}} f_{\mathrm{a}} \delta(\boldsymbol{r} - \boldsymbol{R}_n)\,, \tag{8.1.16}$$

where f_{a} has the dimension of length and is called the *scattering length*. The interaction governing magnetic scattering is proportional to the magnetic moment density.

To determine the scattered beam, we continue to assume that a plane wave of wave vector $\boldsymbol{k}$ is incident upon the sample, and the beam reaching the detector can be considered as a plane wave with wave vector $\boldsymbol{k}'$. The transition probability is proportional to the absolute square of the transition matrix element

$$\langle \boldsymbol{k}|U(\boldsymbol{r})|\boldsymbol{k}'\rangle = \frac{1}{V}\int U(\boldsymbol{r}) \mathrm{e}^{-\mathrm{i}(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{r}}\,\mathrm{d}\boldsymbol{r}\,. \tag{8.1.17}$$

In what follows, $|\boldsymbol{k}\rangle$ does not stand for a plane wave normalized with a factor $1/V^{-1/2}$ but for $\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}$, thus the volume factor $1/V$ is absent from the matrix element.

When the potential is written as a sum of atomic potentials, the above matrix element can also be decomposed into the sum of contributions from individual atoms:

$$\langle \boldsymbol{k}|U(\boldsymbol{r})|\boldsymbol{k}'\rangle = \sum_{\boldsymbol{R}_n} \mathrm{e}^{-\mathrm{i}(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{R}_n} \int U_{\mathrm{a}}(\boldsymbol{r} - \boldsymbol{R}_n) \mathrm{e}^{-\mathrm{i}(\boldsymbol{k}-\boldsymbol{k}')\cdot(\boldsymbol{r}-\boldsymbol{R}_n)}\,\mathrm{d}(\boldsymbol{r} - \boldsymbol{R}_n)\,. \tag{8.1.18}$$

The reader may recognize the $\boldsymbol{k} - \boldsymbol{k}'$ component of the Fourier transform of the atomic potential. Provided that this is identical for each atom, it can be taken out of the matrix element:

$$\langle \boldsymbol{k}|U(\boldsymbol{r})|\boldsymbol{k}'\rangle = U_{\boldsymbol{k}-\boldsymbol{k}'} \sum_{\boldsymbol{R}_n} \mathrm{e}^{-\mathrm{i}(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{R}_n}\,. \tag{8.1.19}$$

The second factor on the right-hand side,

$$\sum_{\boldsymbol{R}_n} \mathrm{e}^{-\mathrm{i}(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{R}_n} \tag{8.1.20}$$

is determined by the spatial distribution of the atoms. It turns out to be the Fourier transform of the atomic density function:

$$\sum_{\boldsymbol{R}_n} \mathrm{e}^{-\mathrm{i}(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{R}_n} = \int \rho(\boldsymbol{r})\mathrm{e}^{-\mathrm{i}(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{r}}\,\mathrm{d}\boldsymbol{r}\,, \tag{8.1.21}$$

where

$$\rho(\boldsymbol{r}) = \sum_{\boldsymbol{R}_n} \delta(\boldsymbol{r}-\boldsymbol{R}_n)\,. \tag{8.1.22}$$

Introduction of the scattering vector $\boldsymbol{K} = \boldsymbol{k} - \boldsymbol{k}'$ leads to

$$\rho_{\boldsymbol{K}} = \sum_{\boldsymbol{R}_n} \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{R}_n}\,; \tag{8.1.23}$$

this quantity is called the *structure amplitude.*[4]

Note that up to this point we have not made use of the periodicity of the atomic arrangement, only the fact that the interaction of radiation with an atom is described by a common potential for each atom. Therefore the above definition of the structure amplitude – the Fourier transform of atomic density – can be extended straightforwardly to noncrystalline materials.

Next, consider a crystal with p atoms arranged identically in each primitive cell. Let $\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_p$ denote the atomic positions within a cell. The potential of the entire lattice is obtained by double summation: over the lattice vectors $\boldsymbol{R}_n$ specifying the primitive cell and over the position vectors $\boldsymbol{r}_j$ of each atom within the cell,

$$U(\boldsymbol{r}) = \sum_{\boldsymbol{R}_n}\sum_{\boldsymbol{r}_j} U_j(\boldsymbol{r}-\boldsymbol{R}_n-\boldsymbol{r}_j)\,. \tag{8.1.24}$$

Naturally, one must allow for different potentials for each atom of the basis. The matrix element is then

$$\langle\boldsymbol{k}|U(\boldsymbol{r})|\boldsymbol{k}'\rangle = \sum_{\boldsymbol{R}_n}\sum_{\boldsymbol{r}_j} U_{j,\boldsymbol{K}}\mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot(\boldsymbol{R}_n+\boldsymbol{r}_j)}\,, \tag{8.1.25}$$

where

$$U_{j,\boldsymbol{K}} = \int U_j(\boldsymbol{r})\mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}}\mathrm{d}\boldsymbol{r} \tag{8.1.26}$$

is the Fourier transform of the potential of the jth atom. As mentioned above, for X-ray diffraction the interaction potential is proportional to the density of electrons,

$$U_{j,\boldsymbol{K}} \propto \int \rho_{\mathrm{e}j}(\boldsymbol{r})\mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}}\,\mathrm{d}\boldsymbol{r}\,. \tag{8.1.27}$$

[4] Although many references use the term *structure factor* for this quantity, this name is customarily reserved for its square, $|\rho_{\boldsymbol{K}}|^2$, which appears in the cross section. This is in line with the definition of the same concept on page 18.

This is why the expression on the right-hand side is called the *atomic form factor*; in what follows, it will be denoted by $A_{\boldsymbol{K}}$. For spherically symmetric electron distributions

$$A_{j,K} = \int 4\pi r^2 \rho_{\mathrm{e}j}(r) \frac{\sin Kr}{Kr} \, \mathrm{d}r \,. \tag{8.1.28}$$

For magnetic scattering of neutrons, the density distribution of the electrons responsible for magnetism appears in the so-called *magnetic form factor*.

In crystalline materials, where each primitive cell is decorated with the same basis, the matrix element can be written as the product of two factors,

$$\langle \boldsymbol{k}|U(\boldsymbol{r})|\boldsymbol{k}'\rangle = \sum_{\boldsymbol{R}_n} \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{R}_n} \sum_{\boldsymbol{r}_j} U_{j,\boldsymbol{K}} \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}_j} \,. \tag{8.1.29}$$

By exploiting the proportionality of the potential and the electron density once again, the second factor on the right-hand side may be written as

$$A_{\boldsymbol{K}} = \sum_{\boldsymbol{r}_j} A_{j,\boldsymbol{K}} \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}_j} \,. \tag{8.1.30}$$

This quantity – or its complex conjugate – accounts for the arrangement of the atoms within a cell, and is therefore also called the *structure amplitude*. Note that the structure amplitude is just the Fourier transform of the total electron density of the atoms within the same cell,

$$A_{\boldsymbol{K}} = \int \rho_{\mathrm{e,tot}}(\boldsymbol{r}) \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}} \, \mathrm{d}\boldsymbol{r} \,, \tag{8.1.31}$$

where

$$\rho_{\mathrm{e,tot}}(\boldsymbol{r}) = \sum_j \rho_{\mathrm{e}j}(\boldsymbol{r}) \,. \tag{8.1.32}$$

For nonmagnetic scattering of neutrons $U_{j,\boldsymbol{K}}$ should be replaced by the Fermi pseudopotential in (8.1.25). Since in this case neutrons are scattered by point-like nuclei, the atomic form factor is independent of $\boldsymbol{K}$ and proportional to the nuclear scattering length. The geometrical information contained in the phase factor is therefore weighted by the scattering lengths of individual atoms in the structure amplitude.

8.1.3 Diffraction Cross Section

Having analyzed the matrix element showing up in the transition probability, we may now turn to the determination of the cross section for diffraction. The differential cross section for the scattering of an X-ray beam into the element of solid angle $\mathrm{d}\Omega$ around the direction of $\boldsymbol{k}'$ is obtained directly from the integration of the double differential formulas given in Appendix E with respect to the energy variable:

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \left(\frac{nk}{2\pi\hbar c}\right)^2 \left|\langle \boldsymbol{k}'|\langle f|U(\boldsymbol{r})|i\rangle|\boldsymbol{k}\rangle\right|^2, \tag{8.1.33}$$

where n is the index of refraction of the sample, while $|i\rangle$ and $|f\rangle$ are the initial and final states of the scattering system. The same formula for neutron scattering is

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \left(\frac{m_\mathrm{n}}{2\pi\hbar^2}\right)^2 \left|\langle \boldsymbol{k}'|\langle f|U(\boldsymbol{r})|i\rangle|\boldsymbol{k}\rangle\right|^2. \tag{8.1.34}$$

Scattering length was introduced in the context of neutron scattering on page 247. By analogy, the scattering amplitude $f(\boldsymbol{r})$ defined through

$$U_\mathrm{a}(\boldsymbol{r}) = -\frac{2\pi\hbar c}{nk} f(\boldsymbol{r}) \tag{8.1.35}$$

is often used instead of the atomic potential for other types of scattering as well. As $f(\boldsymbol{r})$ has dimensions of length, it is often called the scattering length, too. Thus for X-ray and neutron diffraction alike, the cross section of elastic scattering by a single atom is the absolute square of the Fourier transform $f_{\boldsymbol{K}}$ of the scattering amplitude:

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = |f_{\boldsymbol{K}}|^2. \tag{8.1.36}$$

If atoms in a bulk material all have the same scattering amplitude and are immobile at positions $\boldsymbol{R}_m$, the cross section for diffraction is

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = |f_{\boldsymbol{K}}|^2 \sum_{\boldsymbol{R}_m,\boldsymbol{R}_n} \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot(\boldsymbol{R}_m-\boldsymbol{R}_n)}. \tag{8.1.37}$$

Note that the second structure-dependent factor on the right-hand side is – apart from a factor of $1/N$ – the Fourier transform of $\Gamma(\boldsymbol{r})$ introduced in (2.1.16), which is related to the pair-correlation function, i.e., the expectation value of the product of two densities. Separation of the $\boldsymbol{K} = 0$ term from $\Gamma(\boldsymbol{K})$ left us with a quantity that we called the structure factor. Thus the cross section for diffraction is proportional to the structure factor provided that the scattering amplitude is the same for each atom. It should be stressed that the validity of this result is not limited to crystalline materials. In crystals, however, the expression given in (8.1.37) is obtained even when the primitive cell contains several atoms, although in this case $f_{\boldsymbol{K}}$ is replaced by a weighted sum of atomic scattering amplitudes,

$$\sum_{\boldsymbol{r}_j} f_{j,\boldsymbol{K}} \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}_j}. \tag{8.1.38}$$

The directional distribution of the scattered beam is determined primarily by the second factor in (8.1.37), the structure factor. For diffraction in crystals this leads to the condition that $\boldsymbol{K} = \boldsymbol{k} - \boldsymbol{k}'$ should be equal to a reciprocal-lattice vector, as

$$\sum_{\boldsymbol{R}_m,\boldsymbol{R}_n} \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot(\boldsymbol{R}_m-\boldsymbol{R}_n)} = N^2 \sum_{\boldsymbol{G}} \delta_{\boldsymbol{K},\boldsymbol{G}}\,, \tag{8.1.39}$$

where we have used the relation (C.1.39). Indeed, diffraction peaks appear only for scattering processes that satisfy the Laue (Bragg) condition, however the intensity of such Bragg peaks is determined by $|f_{\boldsymbol{K}}|^2$, which depends on atomic scattering amplitudes as well as the geometry of the basis.

Next, we shall demonstrate that the result for the cross section of diffraction obtained by the above calculation,

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \Big| \sum_{\boldsymbol{R}_n} \sum_{\boldsymbol{r}_j} f_{j,\boldsymbol{K}} \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot(\boldsymbol{R}_n+\boldsymbol{r}_j)} \Big|^2\,, \tag{8.1.40}$$

does not fundamentally rely on the assumption that the scattered beam is a plane wave: the same result is recovered using the more natural assumption that scattered waves are spherical.

If the scattering center is at the origin then the wavefunction of the scattered particle and the unscattered part of the beam is given by

$$\psi(\boldsymbol{r}) = \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} + f(\Omega)\frac{\mathrm{e}^{\mathrm{i}kr}}{r}\,, \tag{8.1.41}$$

where Ω is the solid angle defined by the direction of $\boldsymbol{r}$. Theoretical considerations show that the prefactor of the second term, $f(\Omega)$ is just the scattering amplitude in that direction, since the scattering cross section for a beam given in the form (8.1.41) is just

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = |f(\Omega)|^2\,. \tag{8.1.42}$$

If the scatterer is not at the origin but at $\boldsymbol{R}_n+\boldsymbol{r}_j$ and the scattering amplitude is f_j then an extra phase factor appears in the wavefunction. As the beam travels only a distance $|\boldsymbol{r}-\boldsymbol{R}_n-\boldsymbol{r}_j|$ until it reaches the detector,

$$\psi = \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{R}_n+\boldsymbol{r}_j)}\left[\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{r}-\boldsymbol{R}_n-\boldsymbol{r}_j)} + f_j(K)\frac{\mathrm{e}^{\mathrm{i}k|\boldsymbol{r}-\boldsymbol{R}_n-\boldsymbol{r}_j|}}{|\boldsymbol{r}-\boldsymbol{R}_n-\boldsymbol{r}_j|}\right]. \tag{8.1.43}$$

Since the size of sample is small compared to the distance of the detector,

$$k|\boldsymbol{r}-\boldsymbol{R}_n-\boldsymbol{r}_j| \approx kr - k\frac{\boldsymbol{r}\cdot(\boldsymbol{R}_n+\boldsymbol{r}_j)}{r} = kr - \boldsymbol{k}'\cdot(\boldsymbol{R}_n+\boldsymbol{r}_j)\,, \tag{8.1.44}$$

as $\boldsymbol{k}'$ points in the direction of the detector located at $\boldsymbol{r}$, and its magnitude is equal to that of $\boldsymbol{k}$. Thus the sum of the incoming plane wave and the outgoing spherical wave takes the form

$$\psi = \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} + f_j(\Omega)\frac{\mathrm{e}^{\mathrm{i}kr+\mathrm{i}\boldsymbol{K}\cdot(\boldsymbol{R}_n+\boldsymbol{r}_j)}}{r}\,. \tag{8.1.45}$$

For several scatterers we have

$$\psi = \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} + \sum_{\boldsymbol{R}_n}\sum_{\boldsymbol{r}_j} f_j(\Omega)\frac{\mathrm{e}^{\mathrm{i}kr+\mathrm{i}\boldsymbol{K}\cdot(\boldsymbol{R}_n+\boldsymbol{r}_j)}}{r}\,, \tag{8.1.46}$$

so the total scattering amplitude is given by

$$f(\Omega) = \sum_{\boldsymbol{R}_n}\sum_{\boldsymbol{r}_j} f_j(\Omega)\mathrm{e}^{\mathrm{i}\boldsymbol{K}\cdot(\boldsymbol{R}_n+\boldsymbol{r}_j)}\,. \tag{8.1.47}$$

When this is substituted into the cross-section formula (8.1.42), expression (8.1.40) is recovered.

Barring the case of large ideal crystals, when a beam of wave number $\boldsymbol{k}$ is incident upon a sample, the intensity of the scattered beam of wave number $\boldsymbol{k}'$ will be shown to be proportional to $|f_{\boldsymbol{K}}|^2$. To determine the crystal structure, the distribution of the diffracted beam and the relative intensity of the Bragg peaks are measured, and then the arrangement of atoms is deduced. When X-ray or electron diffraction is used, the interaction potential is proportional to the electron density – and then their spatial distribution can also be determined.

According to (C.1.36), lattice-periodic charge distributions can be expanded into Fourier series using vectors of the reciprocal lattice,

$$\rho_{\mathrm{tot}}(\boldsymbol{r}) \propto \frac{1}{v}\sum_{\boldsymbol{G}} f_{\boldsymbol{G}}\mathrm{e}^{\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{r}}\,, \tag{8.1.48}$$

where v is the volume of the primitive cell. This implies that if the structure amplitudes were known for each reciprocal-lattice vector, charge density could, in principle, be obtained by an inverse Fourier transformation. The evaluation of diffraction measurements is encumbered by the fact that the structure factor $|f_{\boldsymbol{K}}|^2$ rather than the structure amplitude is amenable to direct measurement, and thus the information contained in the complex phase of the structure amplitude is lost. To overcome this problem, specific assumptions are made about the structure, and iterative methods are used until an atomic configuration is found that is in agreement with measurements. Even then the structure cannot be reconstructed without some ambiguity. In recent years successful attempts have been made to extend holography to the X-ray region, opening the way to X-ray holograms that contain phase information, and provide a three-dimensional image of the neighborhood of the atom.[5]

8.1.4 The Shape and Intensity of Diffraction Peaks

The above considerations are suitable for the determination of the direction of diffraction peaks for infinitely large samples. For finite samples interference

[5] Outstanding contributions to X-ray holography have been made by G. Faigel and M. Tegze. However, there is still a long way to go before practical applications are available.

is expected to remain constructive, and thus the scattered intensity will be maximal in these directions, however, in other directions cancellation will be only partial, leading to small but nonzero intensities. In analogy to optics, one expects that by increasing the size of the sample (that is, the number of scatterers), the diffraction peak becomes sharper, and practically complete cancellation is observed in any direction that does not satisfy the Bragg condition. To determine the angular dependence of intensity, we shall return to the formulation asserting that the transition probability is proportional to the absolute square of the transition matrix element

$$\langle \boldsymbol{k}|U(\boldsymbol{r})|\boldsymbol{k}'\rangle = \int U(\boldsymbol{r})\mathrm{e}^{-\mathrm{i}(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{r}}\,\mathrm{d}\boldsymbol{r}\,, \tag{8.1.49}$$

where $U(\boldsymbol{r})$ is the full potential. In terms of the scattering amplitude $f(\boldsymbol{r})$ instead of the potential, the amplitude of the scattered beam is proportional to

$$f_{\boldsymbol{K}} = \int f(\boldsymbol{r})\mathrm{e}^{-\mathrm{i}(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{r}}\,\mathrm{d}\boldsymbol{r}\,. \tag{8.1.50}$$

In crystalline material atoms are arranged in a periodic array, thus the scattering amplitude is also periodic and satisfies condition (5.1.2). It is therefore possible to expand it into a Fourier series in terms of the reciprocal-lattice vectors. Using the form

$$f(\boldsymbol{r}) = \frac{1}{v}\sum_{\boldsymbol{G}} f_{\boldsymbol{G}}\mathrm{e}^{\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{r}}\,, \tag{8.1.51}$$

where v is the volume of the primitive cell, the amplitude of the scattered beam is found to be proportional to

$$f_{\boldsymbol{K}} = \frac{1}{v}\sum_{\boldsymbol{G}} f_{\boldsymbol{G}}\int \mathrm{e}^{-\mathrm{i}(\boldsymbol{k}-\boldsymbol{k}'-\boldsymbol{G})\cdot\boldsymbol{r}}\,\mathrm{d}\boldsymbol{r}\,. \tag{8.1.52}$$

For infinite crystals scattered beams arise only in the directions that satisfy the condition $\boldsymbol{k}' = \boldsymbol{k} - \boldsymbol{G}$; cancellation is perfect in all other directions.

In finite crystals, an arbitrary position $\boldsymbol{r}$ is written as the vector sum of the lattice point $\boldsymbol{R}_n$ associated with the primitive cell and the position vector $\boldsymbol{u}$ within the cell:

$$\boldsymbol{r} = \boldsymbol{R}_n + \boldsymbol{u} = n_1\boldsymbol{a}_1 + n_2\boldsymbol{a}_2 + n_3\boldsymbol{a}_3 + u_1\boldsymbol{a}_1 + u_2\boldsymbol{a}_2 + u_3\boldsymbol{a}_3\,. \tag{8.1.53}$$

The coordinates u_i of the position within the primitive cell satisfy the condition

$$0 \le u_j \le 1\,, \qquad j = 1, 2, 3. \tag{8.1.54}$$

Next, the vector $\boldsymbol{K} = \boldsymbol{k} - \boldsymbol{k}'$ is expressed in terms of reciprocal-lattice vectors. From the foregoing it is intuitively obvious that scattering is maximal if $\boldsymbol{K}$ is precisely equal to a vector of the reciprocal lattice. Since we are interested

in the shape of the peak, $\boldsymbol{K}$ is now allowed to differ by an amount $\boldsymbol{q}$ from a reciprocal-lattice vector $\boldsymbol{G}$. Expressed in terms of the primitive vectors $\boldsymbol{b}_1$, $\boldsymbol{b}_2$, and $\boldsymbol{b}_3$ of the reciprocal lattice,

$$\boldsymbol{K} = \boldsymbol{G} + \boldsymbol{q} = h\boldsymbol{b}_1 + k\boldsymbol{b}_2 + l\boldsymbol{b}_3 + q_1\boldsymbol{b}_1 + q_2\boldsymbol{b}_2 + q_3\boldsymbol{b}_3\,. \tag{8.1.55}$$

As the scattered beam is confined to a small cone around the direction of $\boldsymbol{k}' = \boldsymbol{k} - \boldsymbol{G}$, the fractional numbers q_i are chosen to fall in the interval

$$-\tfrac{1}{2} \le q_j \le \tfrac{1}{2}\,, \qquad j = 1, 2, 3. \tag{8.1.56}$$

To evaluate the integral in (8.1.52) for a fixed $\boldsymbol{q}$, it is separated into an integral over the primitive cell and a sum over primitive cells:

$$f_{\boldsymbol{q}} = \frac{1}{v} f_{\boldsymbol{G}} \sum_{\boldsymbol{R}_n} \int_v \mathrm{e}^{-\mathrm{i}\boldsymbol{q}\cdot(\boldsymbol{R}_n + \boldsymbol{u})}\,\mathrm{d}\boldsymbol{u}\,. \tag{8.1.57}$$

Suppose that the shape of the sample is such that there are N_1, N_2, and N_3 primitive cells along the directions of the primitive vectors $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, and $\boldsymbol{a}_3$. As the scalar product of a direct- and a reciprocal-lattice vector satisfies (5.2.12),

$$f_{\boldsymbol{q}} = f_{\boldsymbol{G}} M_1 M_2 M_3\,, \tag{8.1.58}$$

where

$$M_j = \sum_{n_j=0}^{N_j-1} \int_0^1 \mathrm{e}^{-2\pi\mathrm{i}(n_j+u_j)q_j}\,\mathrm{d}u_j\,. \tag{8.1.59}$$

Summation and integration can be separated. Each operation is then elementary, leading to

$$M_j = S_j M_j'\,, \tag{8.1.60}$$

where

$$S_j = \sum_{n_j=0}^{N_j-1} \mathrm{e}^{-2\pi\mathrm{i}n_j q_j} = \frac{\mathrm{e}^{-2\pi\mathrm{i}N_j q_j} - 1}{\mathrm{e}^{-2\pi\mathrm{i}q_j} - 1}\,, \tag{8.1.61}$$

and

$$M_j' = \int_0^1 \mathrm{e}^{-2\pi\mathrm{i}u_j q_j}\,\mathrm{d}u_j = \frac{1 - \mathrm{e}^{-2\pi\mathrm{i}q_j}}{2\pi\mathrm{i}q_j}\,, \tag{8.1.62}$$

that is

$$M_j = \frac{1 - \mathrm{e}^{-2\pi\mathrm{i}N_j q_j}}{2\pi\mathrm{i}q_j}\,. \tag{8.1.63}$$

The intensity I of the scattered beam contains the absolute square of the amplitude. Using

$$\left|\frac{1 - \mathrm{e}^{-2\pi\mathrm{i}N_j q_j}}{2\pi\mathrm{i}q_j}\right|^2 = \frac{\sin^2 \pi q_j N_j}{\pi^2 q_j^2}\,, \tag{8.1.64}$$

the intensity distribution is given by

$$I(\boldsymbol{q}) = |f_{\boldsymbol{G}}|^2 \, \frac{1}{v^2} \, \frac{\sin^2 \pi q_1 N_1}{\pi^2 q_1^2} \, \frac{\sin^2 \pi q_2 N_2}{\pi^2 q_2^2} \, \frac{\sin^2 \pi q_3 N_3}{\pi^2 q_3^2} \tag{8.1.65}$$

for the diffraction peak associated with the reciprocal-lattice vector $\boldsymbol{G}$ and the direction characterized by the components q_i. This function has a sharp maximum at $\boldsymbol{q} = 0$, where its value is proportional to the square of the number of scatterers – that is, the square of the sample volume. This result is counterintuitive: the intensity of the scattered beam is expected to be proportional to the volume of the sample (and not its square). If we were to determine not just the peak height but the total scattered intensity, it would be proportional to the volume of the sample, as the width of peaks decreases inversely with the volume of the sample. Nevertheless this result indicates the necessity of a more rigorous discussion of the rays scattered from the sample interior. We shall revisit this point later, when the dynamical theory of diffraction has been outlined.

8.1.5 Cancellation in Structures with a Polyatomic Basis

In structures with a monatomic basis, where each primitive cell contains a single atom, the intensity of scattering associated with vector $\boldsymbol{G}$ of the reciprocal lattice is simply proportional to $|f_{\boldsymbol{G}}|^2$. The situation is different in crystals with a basis of several atoms. Here intensity relations depend on the geometry of the atoms of the basis as well as the magnitude of the individual scattering amplitudes. The situation is substantially simplified when the atoms of the basis are all identical. The atomic scattering factor can then be factored out, leaving behind a structure factor that conveys information about the structure alone. Consider, for example, a diamond lattice, which is a face-centered cubic structure with a diatomic basis. In terms of the edge vectors of the face-centered cubic Bravais cell the coordinates are 000 and $\frac{1}{4}\frac{1}{4}\frac{1}{4}$. Using the primitive vectors given in (7.2.9) leads to the same coordinates: the second carbon atom is at $\frac{1}{4}(\boldsymbol{a}_1 + \boldsymbol{a}_2 + \boldsymbol{a}_3)$. The structure amplitude for reciprocal-lattice vector $\boldsymbol{G} = h\boldsymbol{b}_1 + k\boldsymbol{b}_2 + l\boldsymbol{b}_3$ – where $\boldsymbol{b}_i$ is defined by (7.2.12) – is

$$A_{hkl} = 1 + \mathrm{e}^{-\mathrm{i}(h\boldsymbol{b}_1 + k\boldsymbol{b}_2 + l\boldsymbol{b}_3)\cdot(\boldsymbol{a}_1 + \boldsymbol{a}_2 + \boldsymbol{a}_3)/4} = 1 + \mathrm{e}^{-\mathrm{i}\pi(h+k+l)/2} \, . \tag{8.1.66}$$

Depending on the indices hkl of the Bragg peak, this quantity takes the values $2, 0$, or $1 \pm \mathrm{i}$. The structure amplitude vanishes when $h + k + l$ is even but not divisible by four ($h + k + l = 4j + 2$). This means that there are directions in which Bragg peaks are absent even though the Laue (Bragg) condition is satisfied by the reciprocal-lattice vector $\boldsymbol{G}_{hkl}$. For other values of hkl there are Bragg peaks, however, when $h+k+l$ is odd ($h+k+l = 2j+1$), their intensity is just half of those peaks for which $h+k+l$ is divisible by four ($h+k+l = 4j$). As Fig. 8.5 shows, those lattice points of the bcc reciprocal lattice for which

the structure amplitudes – and hence the Bragg peak intensities – are equal are located in planes perpendicular to the space diagonal.

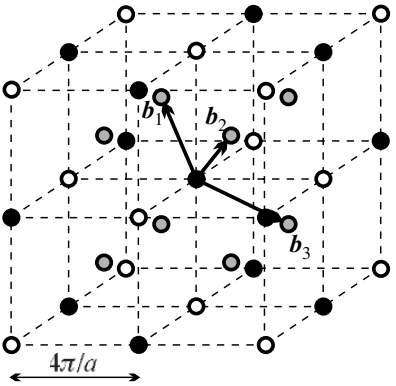


Fig. 8.5. Reciprocal-lattice vectors of the diamond structure that give rise to diffraction peaks of different intensities. The structure amplitude for vectors pointing to full circles, shaded circles, and empty circles are 2, $1 \pm \mathrm{i}$, and 0. The intensity ratio of the corresponding Bragg peaks is therefore $4 : 2 : 0$

The reason for this is that rays scattered by the two carbon atoms of the basis may interfere constructively or destructively; they may even cancel out perfectly. Observation of these cancellations permits experimentalists to distinguish the diamond structure from face-centered structures with a monatomic basis.

When the basis contains two different atoms – as in the sphalerite structure, where Zn and S atoms are at points 000 and $\frac{1}{4}\frac{1}{4}\frac{1}{4}$, respectively –, the scattering amplitude of the primitive cell is calculated from atomic scattering amplitudes using weights characteristic of the structure:

$$f_{hkl} = f_{\mathrm{Zn}} + f_{\mathrm{S}} \mathrm{e}^{-\mathrm{i}\pi(h+k+l)/2} \,. \tag{8.1.67}$$

This shows that there are no reciprocal-lattice vectors for which cancellation is perfect. From the intensity distribution of the Bragg peaks one may infer the atomic scattering factors and atomic positions within the cell – that is, the nature of the crystal.

Similarly to the case of the diamond structure, a cancellation (albeit an apparent one) is observed when the crystal has a centered Bravais lattice – but this information is not known prior to the evaluation of the diffraction measurements. As a simple example consider a face-centered cubic lattice with a lattice constant a. When, based on macroscopic properties, the crystal structure is assumed to possess cubic symmetry, one may attempt to interpret the diffraction peaks in terms of the reciprocal lattice of a simple cubic crystal. Using the edge vectors $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$ of the Bravais cell (see Fig. 8.6(a)), the

coordinates of the four atoms of the cell are 000, $\frac{1}{2}\frac{1}{2}0$, $\frac{1}{2}0\frac{1}{2}$, and $0\frac{1}{2}\frac{1}{2}$. The counterparts of the direct-lattice edge vectors $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ are the reciprocal-space vectors $\boldsymbol{a}^*$, $\boldsymbol{b}^*$, and $\boldsymbol{c}^*$ of length $2\pi/a$. In terms of these the structure amplitude for the vector $\boldsymbol{G} = h\boldsymbol{a}^* + k\boldsymbol{b}^* + l\boldsymbol{c}^*$ is

$$A_{hkl} = \sum_j \mathrm{e}^{-2\pi\mathrm{i}(hx_j+ky_j+lz_j)} = 1+\mathrm{e}^{-\mathrm{i}\pi(h+k)}+\mathrm{e}^{-\mathrm{i}\pi(h+l)}+\mathrm{e}^{-\mathrm{i}\pi(k+l)}\,. \quad (8.1.68)$$

The value of this quantity is 4 when the Miller indices hkl are all even or all odd. When odd as well as even indices occur, the structure factor vanishes – and so no scattered beam emerges in the corresponding direction. Certain Bragg peaks are absent – while they should be present if the crystal had a simple cubic structure. Empty circles in Fig. 8.6(b) indicate those vectors $\boldsymbol{G}$ of the reciprocal lattice for which the structure factor vanishes, and full circles those for which the Bragg peak is finite.

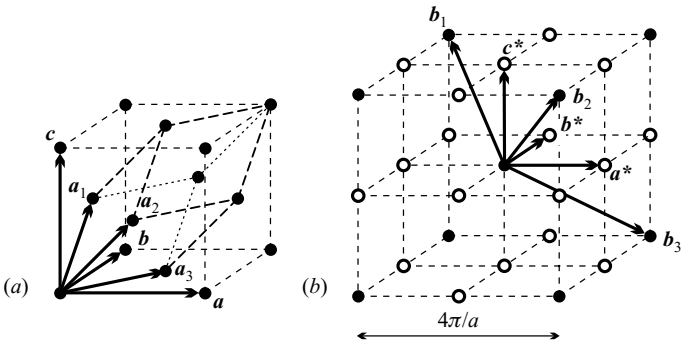


Fig. 8.6. (a) The primitive cell and the Bravais cell for a face-centered cubic crystal. (b) The body-centered cubic reciprocal lattice of the face-centered cubic direct lattice is indicated by full circles. The simple cubic lattice with the additional lattice points (empty circles) is the reciprocal of the lattice spanned by the primitive vectors $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$

Compared to the simple cubic lattice, certain Bragg peaks are thus absent in the corresponding face-centered cubic crystal. This can be interpreted as the result of destructive interference between the rays scattered by the atoms at the vertices and face centers. This is illustrated in Fig. 8.7. The planes drawn across the atoms at the vertices of the simple cubic lattice are indicated by solid lines, while the extra planes appearing in an fcc lattice due to centering by dashed lines.

The figure shows that the spacing between the atomic planes of an fcc crystal is half of that for the simple cubic crystal from which it is obtained by centering. Let us consider directions for which the phase difference between rays reflected from adjacent planes of the simple cubic lattice is an odd multiple of 2π – i.e., the path difference is an odd multiple of λ. Among themselves,

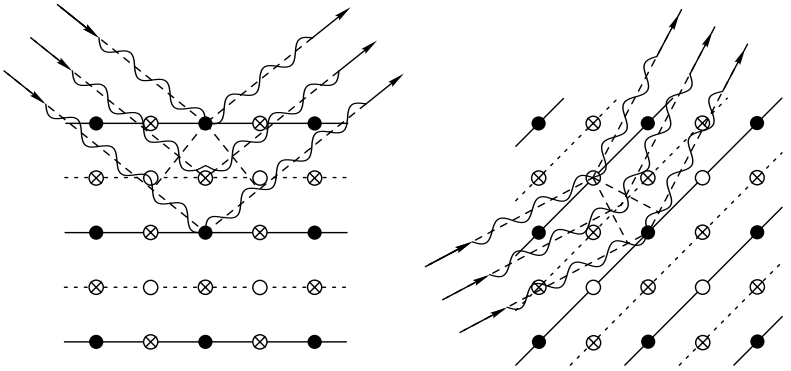


Fig. 8.7. Projection of the face-centered cubic crystal onto the plane perpendicular to the [001] direction, with the (100) and (110) planes. Full circles, empty circles and ⊗ symbols indicate atoms at vertices, base centers, and side centers. Solid lines show those members of the family of planes that contain atoms at the vertices of the Bravais cell. Atomic planes drawn with dashed lines are the result of centering

these rays interfere constructively. When, however, the rays scattered from the extra planes of the fcc lattice are also taken into account, cancellation occurs, because these rays are precisely in the opposite phase.

However, this cancellation is only apparent: it occurs only when the Bragg condition is employed for the atomic planes of the simple cubic lattice whose centering gives the real lattice, the face-centered cubic one. Similarly: when using the Laue condition, the absence of scattering for certain vectors of the reciprocal lattice is the consequence of considering a simple rather than a face-centered cubic lattice. The reciprocal lattice should be defined in terms of the primitive vectors of the true primitive cell $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, and $\boldsymbol{a}_3$ rather than those of the Bravais cell (conventional unit cell). Since the (direct-lattice) Bravais cell is larger than the primitive cell, the reciprocal lattice $\boldsymbol{G}^* = h\boldsymbol{a}^* + k\boldsymbol{b}^* + l\boldsymbol{c}^*$ – generated by the vectors $\boldsymbol{a}^*$, $\boldsymbol{b}^*$, $\boldsymbol{c}^*$ associated with the edge vectors of the Bravais cell – is denser than the true reciprocal lattice, spanned by the primitive vectors (7.2.12). As we have seen, the latter is a body-centered cubic lattice with edge length $4\pi/a$. This lattice is shown by the full circles in Fig. 8.6. Scattered beams emerge only in those directions for which the Laue condition $\boldsymbol{k} - \boldsymbol{k}' = \boldsymbol{G}$ is satisfied by the vectors of reciprocal lattice of the face-centered cubic lattice.

8.1.6 The Dynamical Theory of Diffraction

When calculating the intensity of the diffraction peak it was assumed that the atoms of the sample scatter the coherent incident beam independently of each other, i.e., a photon scattered by an atom does not undergo another scattering. Neglecting multiple scattering processes may be justified for small samples

and powders of polycrystalline materials, where interference is assumed to be absent for beams scattered by grains of different orientation. However, for larger single crystals it is no longer true that rays scattered by atoms deep inside the sample and close to its surface are equally intense. A beam incident upon the sample is scattered by the topmost atomic layer, giving a reflected and a transmitted beam. The latter arrives at the second layer of atoms, which reflects a part of it and transmits another to the third layer. Naturally, radiation reflected from deeper layers may bounce back from upper ones. This is illustrated in Fig. 8.8.

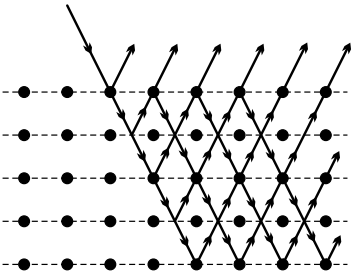


Fig. 8.8. Interference of the waves reflected from and transmitted by subsequent layers of the crystal

Beams undergoing multiple scattering in the sample may interfere with each other, building up an internal field. To determine this field the phase shift of the waves occurring in scattering processes must not be ignored. The theory that takes into account the amplitude and phase relationships of beams scattered in the sample is called the *dynamical theory of diffraction*,[6] as opposed to the *kinematical theory* discussed above. Without going into details, we shall briefly present the crucial elements of the theoretical description, pointing out the differences with the problem discussed in Chapter 25, the scattering of light by solids (where Fresnel's equations are used).

The electromagnetic field built up inside the sample obeys Maxwell's equations. Since conductivity practically vanishes for X-ray frequencies, these simplify to

$$\operatorname{curl} \boldsymbol{H} = \frac{\partial \boldsymbol{D}}{\partial t}\,, \qquad \operatorname{curl} \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t}\,. \tag{8.1.69}$$

For nonmagnetic materials $\boldsymbol{B} = \mu_0 \boldsymbol{H}$ to a good approximation, while the dielectric constant relating $\boldsymbol{D}$ and $\boldsymbol{E}$ is assumed to be lattice periodic within the sample,[7] and can thus be expanded into a Fourier series in terms of the reciprocal-lattice vectors:

[6] C. G. Darwin (1914) and P. P. Ewald (1916).

[7] This assumption is not needed in the optical region where the dielectric constant is assumed to be uniform both inside and outside the sample.

$$\epsilon_{\mathrm{r}} = \frac{1}{v} \sum_{\boldsymbol{G}} \epsilon_{\boldsymbol{G}} \mathrm{e}^{\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{r}} \,. \tag{8.1.70}$$

Let a plane wave of wave vector $\boldsymbol{k}$ be incident on the sample and scattered by the periodic structure. The scattered wave is a mixture of waves with wave vectors $\boldsymbol{k}_{\boldsymbol{G}} = \boldsymbol{k} + \boldsymbol{G}$. The interior electric and magnetic fields are therefore written as the linear combinations of plane waves with such wave vectors. Substituting them back into Maxwell's equations gives relationships among the Fourier coefficients. The problem is substantially simplified by the assumption that the Bragg condition is satisfied for a single reciprocal-lattice vector, i.e., the amplitude is important for a single component $\boldsymbol{k} + \boldsymbol{G}$ apart from $\boldsymbol{k}$.

The next important step is to satisfy the matching conditions across the boundary surface. In contrast to the optical region, where the directions and amplitudes of the incident, reflected and transmitted beams need to be matched, the beam reflected from the surface is practically absent in the X-ray region as the difference of ϵ_{r} from unity is tiny (on the order of 10^{-4} or less), and so the amplitude of the incident beam is equal to that of the transmitted beam at the surface. Only wave vectors need to be matched; for this it should be borne in mind that inside the sample $\boldsymbol{k}$ may be complex because of absorption.

Calculations show that if the sample is sufficiently thick and the crystal is ideal – that is, subsequent layers are perfectly parallel – then cancellation is exact in the transmitted beam, and so the diffracted beam emerges in a very narrow pencil – that is less than an arc minute across, yet finite in diameter – around the direction satisfying the Bragg condition. The width of the beam is determined by the size of the sample and the Fourier transform $f_{\boldsymbol{G}}$ of the spatial distribution of scatterers. The total intensity of the Bragg peak is proportional to the volume, just like in the kinematical theory, however, here it is proportional to $|f_{\boldsymbol{G}}|$ instead of its square. Of course, intensity also depends on whether the Laue or Bragg case is considered. The former corresponds to a diffracted beam emerging on the other side of the sample – and thus attenuated by absorption –, while the latter corresponds to a diffracted beam emerging from the same surface on which the incoming beam is incident.

To evaluate the intensity of the scattered beam one has to take into account that atoms of the crystal are not fixed rigidly at their equilibrium positions but oscillate around it. The amplitude of this vibration becomes larger at higher temperatures, and its consequences cannot be ignored in a precise study of diffraction. We shall see in Chapter 12 on lattice vibrations that the Bragg condition is still valid for elastic scattering: even when atoms oscillate harmonically, diffraction peaks continue to be sharp, however the intensity of the scattered beam decreases with increasing temperature.

8.2 Experimental Study of Diffraction

The previous introduction into the theory of diffraction has shown that structural information can be extracted from reciprocal-lattice vectors and the intensity of Bragg peaks by determining the directions in which elastically scattered beams emerge when a beam of known wavelength is shone on a sample from a well-defined direction. In an ideal experimental setup the source provides a well-collimated beam of monoenergetic (monochromatic) particles, that is, the wave vector $\boldsymbol{k}$ is the same for each particle. When the beam is not monochromatic, a single-crystal monochromator can be inserted so that only a particular wavelength λ that satisfies the Bragg condition is reflected. The wavelength of the monochromatized beam can be changed by rotating the monochromator.

The number of particles scattered elastically in a certain direction is then measured by suitable detectors. For a particular value of λ, the Bragg peak can be found by varying the angle 2θ of the detector. In the case of X-ray diffraction the intensity of the scattered beam is often detected using a film instead of a photon counter. The schematic setup of the measurement is shown in Fig. 8.9.

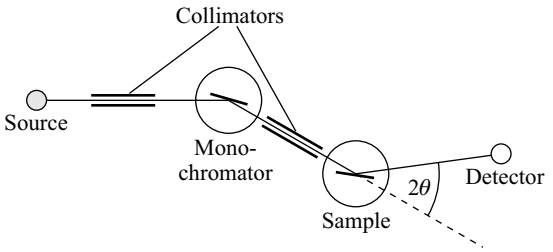


Fig. 8.9. Schematic setup of diffraction measurements

8.2.1 Characteristic Properties of Different Types of Radiation

Let us now examine what photon, neutron, and electron energies provide the most effective wavelength region for structural analysis. We shall start off with the relationship between the photon energy and wavelength, $\varepsilon = cp = hc/\lambda$. When photon energies are measured in units of keV and wavelengths in angstroms, the values listed in Appendix A for physical constants give

$$\lambda\,[\text{Å}] = \frac{12.4}{\varepsilon\,[\text{keV}]}\,. \tag{8.2.1}$$

To obtain an X-ray beam with a wavelength of 1 Å, photon energies on the order of 10 keV are necessary.

Photons emitted in a synchrotron by accelerated electrons provide intense and collimated beams with a wide energy spectrum. Then a photon beam propagating in a well-defined direction is selected, in which photon energies are an order of magnitude larger than in ordinary X-rays, i.e., whose wavelength is on the order of 10^{-1} Å. This provides a much better resolution in momentum space than a traditional X-ray source. Because of the large intensity, diffraction measurements may be performed even on tiny (µg-sized) crystallites; or, short measurement times open the way to studying time-varying structures with a time resolution of about a millisecond.

Neutron energies are given by the relation $\varepsilon = p^2/2m_\mathrm{n}$, where $m_\mathrm{n} = 1.675 \times 10^{-27}\,\mathrm{kg}$ is the neutron mass. The momentum p is related to the de Broglie wavelength λ by $p = h/\lambda$, thus in terms of the energy measured in electronvolts, the wavelength in angstroms is given as

$$\lambda\,[\text{Å}] = \frac{0.28}{(\varepsilon\,[\text{eV}])^{1/2}}\,. \tag{8.2.2}$$

To obtain a wavelength of about 1 Å, the energy has to be $\varepsilon \approx 0.1\,\text{eV}$. This energy is close to the thermal energy at room temperature.[8] Thermal neutrons – neutrons emerging from a fission reactor and then slowed down to thermal energies by the moderator – are thus of the desired wavelength.

Because of their smaller mass, the wavelength–energy relationship for electrons derived from the formula $\varepsilon = p^2/2m_\mathrm{e}$ is

$$\lambda\,[\text{Å}] = \frac{12.247}{(\varepsilon\,[\text{eV}])^{1/2}}. \tag{8.2.3}$$

To achieve a wavelength of 1 Å, electrons need to be accelerated to energies on the order of 100 eV. The experimental method using electrons with energies between 10 and 300 eV (wavelength: 0.7–4 Å) is called *low-energy electron diffraction* (LEED). When the beam energy is higher, we speak of *high-energy electron diffraction* (HEED).

Since diffraction is an elastic process, the energy of the scattered particle is irrelevant. From this point of view there should be no difference in the applicability of the three types of beam. Particle energies become important when we turn to the suitability of each particular type of radiation to the investigation of the internal dynamics of solids (where typical energies range from 10 to 100 meV). The accuracy of measuring changes in energy caused by low-energy excitations is limited by the resolution of the measuring apparatus. The energy 10^{-1} eV of one-angstrom neutrons is ideally suited to studying internal excitations. The amount of information that can be gathered with electrons and photons is much smaller because of the high energy of the particles.

[8] According to the equipartition theorem, the mean kinetic energy of particles at temperature T is $\frac{3}{2}k_\mathrm{B}T$. At room temperature this corresponds to an energy of 38 meV or a wavelength of 1.4 Å.

Even though the only relevant quantity for the Bragg or Laue condition is the wave vector of the incoming radiation (since the energy is irrelevant in elastic scattering), the three radiations are not equally suited to probing a given structure via diffraction. As each radiation interacts differently with the various constituents of the sample, the best choice depends on the particular problem. Photon and electron beams interact with electrons in the solid, thus information is gathered directly about the spatial distribution of electrons. Neutrons are scattered by nuclei, and they interact with the electron cloud, too, if the latter has a magnetic moment. They can thus probe atomic as well as magnetic structures. Neutron diffraction measurements may complement the results obtained via X-ray diffraction even for magnetically disordered systems. The cross section for X-ray scattering depends on the number of electrons in the electron cloud: it goes with the square of the atomic number. The cross section for neutron scattering, on the other hand, is not a monotonic function of the atomic number, and it also depends on the isotopic composition. Depending on the composition of the sample, either type of radiation may be better suited to specifying atomic positions within the primitive cell. For example, unlike X-ray scattering, neutron scattering is particularly adapted for determining the structure of samples that contain hydrogen.

By using a radiation whose wavelength is much smaller than the atomic dimensions, the spatial distribution of electron density may also be mapped out. HEED is the method of choice for this, as the wavelength of 100 keV electrons is $\lambda \approx 0.04$ Å. Another advantage of HEED over LEED is that high-energy electrons can penetrate deep inside the sample, thus this method permits us to study the interior of the sample. Since beams used in LEED usually penetrate only into the topmost layer of a few angstroms, this method is adapted for the study of the structure at or close to the surface.

How much of the radiation is absorbed by the sample; how deeply does it penetrate into the sample, and so to what extent can bulk properties be studied? These questions arise for all other types of radiation as well. Neutrons can penetrate most deeply into the sample. Depending on the material, they may penetrate through a sample several inches thick without absorption. On the other hand, 1 Å X-ray photons typically penetrate into metals only to a depth of 30–100 µm – but even this may be enough to measure bulk properties. Since their penetration depth is much larger in the appropriate wavelength region, synchrotron radiation obviously probes bulk properties.

There is, nevertheless, another typical length that has to be taken into account, the *extinction length*. Its typical value is 10 µm for X-rays, 100 µm for neutrons, and a mere 10^{-2} µm for electrons. For thinner samples the kinematic theory of diffraction may be used without reservations. For thicker ones it can only be used if the crystal is not ideal but features a mosaic structure.[9] Otherwise internal cancellations can no longer be ignored, and thus the dynamical theory of diffraction has to be employed.

[9] Mosaic structures will be presented in the next chapter.

8.2.2 The Ewald Construction

According to conditions (8.1.1) and (8.1.5) for diffraction, when a beam of a specific wavelength λ is incident on a crystalline sample, scattered radiation emerges only in some particular directions. The *Ewald construction*[10] is of great help in visualizing this.

Since the condition for diffraction is most easily expressed in terms of reciprocal-lattice vectors, we shall consider the reciprocal lattice of the putative Bravais lattice of the examined crystal. Figure 8.10 shows its sectional view in the plane of scattering. The incident wave vector $\boldsymbol{k}$ is drawn in such a way that its tip is at a lattice point of the reciprocal lattice. Centered at the starting point of the arrow $\boldsymbol{k}$, a sphere (circle) of radius $k = |\boldsymbol{k}|$ is then drawn. This is the *Ewald sphere*.

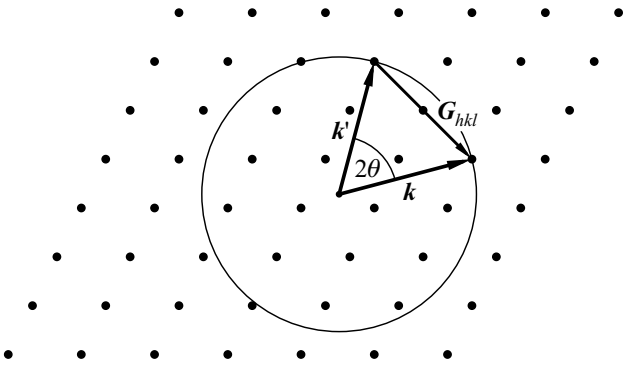


Fig. 8.10. The Ewald construction illustrating the condition for diffraction in the reciprocal lattice

By its construction, the Ewald sphere passes through a point of the reciprocal lattice (lying at the tip of vector $\boldsymbol{k}$). If it happens to pass through other points of the reciprocal lattice then the Laue condition is satisfied by the vectors $\boldsymbol{k}'$ drawn from the center of the sphere to each of these points as $|\boldsymbol{k}| = |\boldsymbol{k}'|$ and $\boldsymbol{k} - \boldsymbol{k}'$ is a vector $\boldsymbol{G}_{hkl}$ of the reciprocal lattice. The Bragg peak with indices hkl is observed in the direction of $\boldsymbol{k}'$.

In LEED the wavelength is much smaller than the lattice constant, and so the magnitude of the wave vector – the radius of the Ewald sphere – is large compared to the lattice constant of the reciprocal lattice. Because of its small curvature, the Ewald sphere may pass very close to a great number of reciprocal-lattice points. When the sample is not too large, many reflections appear close to the incident direction because of the finite width of the diffraction peaks.

[10] P. P. Ewald, 1913.

This is not the case for X-ray and neutron diffraction. When the orientation of the crystal with respect to the incident beam is arbitrary, there may be no scattering at all. To meet the Bragg condition, one has to change either the wavelength of the incident beam by rotating the axis of the single-crystal monochromator; or the relative orientation of the beam and the crystallographic axes by rotating the sample; or the direction of the detected (scattered) beam by moving the particle counter. When the detector is a film or a set of counters that covers a large proportion of the total solid angle 4π, there are still two degrees of freedom in choosing the wavelength of the incident beam and the relative orientation of the beam and the crystal at will. In the most sophisticated method the wavelength is fixed and the complete diffraction pattern is measured for several orientations of the sample. Simpler methods can be employed when the task is not the precise determination of a completely unknown structure but the quick verification of the putative structure or the determination of the lattice parameters for a known structure type.

8.2.3 Diffraction Methods

One possibility is to irradiate a sample with a collimated but not monochromatic beam that contains all the wavelengths from $\lambda_{\min}$ up to $\lambda_{\max}$. This is called the *Laue method*. The direction of the wave vector is now definite, however its magnitude ranges between $k_{\min}$ and $k_{\max}$.[11] If the Ewald spheres were drawn for all possible values of the incoming wavelength, they would fill the region between two spheres of radii $k_{\min}$ and $k_{\max}$, as shown in Fig. 8.11.

The Laue condition is satisfied by every vector of the reciprocal lattice within this region, and so scattering is observed in several directions. Based on the Ewald construction it can be shown that if the orientation of the sample is such that the beam is incident along a high-symmetry direction of the crystal then the Laue pattern also possesses the rotational symmetries of the crystal around the axis. Figure 8.12 shows Laue patterns with the hexagonal structure of beryllium and the trigonal structure of quartz.

Another method for meeting the condition for diffraction is to rotate the crystal around an axis perpendicular to the direction of the incident beam – or to vibrate it covering a sufficiently large angular range. This is the *rotating-crystal method*.

The sample is oriented in such a way that the axis of rotation should be along a crystallographic axis – for example, in the direction of the primitive vector $\boldsymbol{a}_1$ of the crystal lattice. Because of the construction of the reciprocal lattice the plane spanned by two appropriately chosen primitive vectors ($\boldsymbol{b}_2$ and $\boldsymbol{b}_3$ in our case) is perpendicular to the axis of rotation. When the Laue condition is satisfied by a reciprocal-lattice vector in this plane, the propagation direction of the scattered beam also lies in the same plane. To determine

[11] Synchrotrons provide such beams, at much higher intensities than traditional X-ray sources. This is why the method is rapidly gaining popularity.

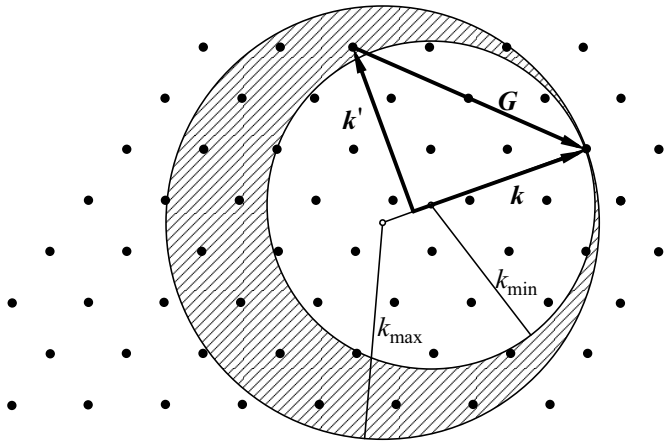


Fig. 8.11. Reciprocal-lattice vectors satisfying the condition for diffraction in the Laue method

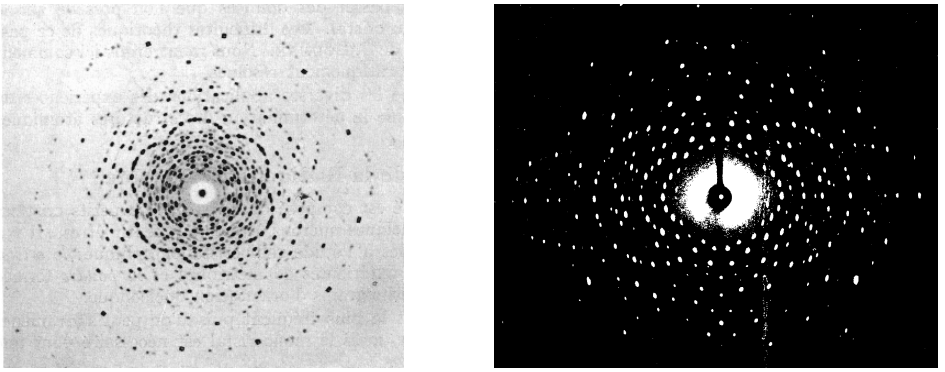


Fig. 8.12. Diffraction patterns along the direction of a crystallographic axis in beryllium and quartz obtained with the Laue method

the directions satisfying the condition for diffraction, Fig. 8.13 shows the reciprocal lattice and the Ewald sphere for a particular orientation of the crystal.

When the crystal is rotated, the reciprocal lattice co-rotates with it. In the Ewald construction this corresponds to the rotation of the reciprocal lattice about the origin. Scattering occurs when a rotated reciprocal-lattice vector is exactly on the surface of the Ewald sphere. Recorded on a film, the scattered radiation appears as a series of dots. Further series of dots may appear above and below this scattering plane. These diffraction peaks correspond to reciprocal-lattice vectors that have a nonvanishing component along the axis of rotation. Such an experimental record is shown in Fig. 8.14.

In the third widely used method measurements are not performed on a single crystal with a definite orientation but on a powder sample that contains

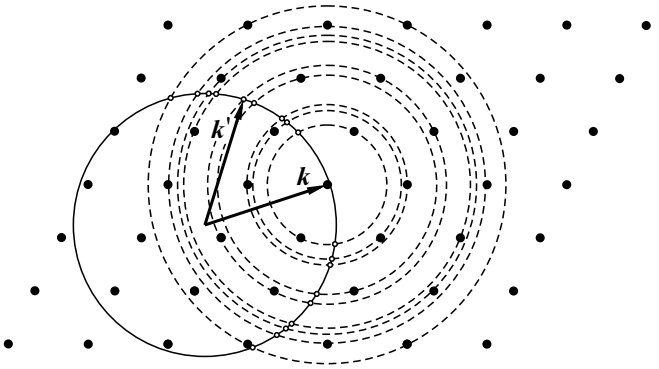


Fig. 8.13. The Ewald construction for a crystal rotating around an axis perpendicular to the direction of the incident beam. The solid circle is the section of the Ewald sphere, while dashed circles show the loci of the tips of reciprocal-lattice vectors when the sample is rotated

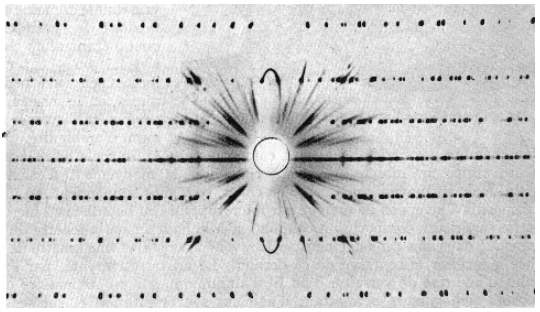


Fig. 8.14. Diffraction pattern of a quartz crystal obtained with the rotating-crystal method

grains with all possible orientations. The *powder method* was first developed by P. Debye[12] and P. Scherrer in 1916, hence it is also called the *Debye–Scherrer method*. To interpret the diffraction pattern assume that the Laue condition is met for a certain grain along a direction $\boldsymbol{k}'$ by the reciprocal-lattice vector $\boldsymbol{G}$. Since the same vector $\boldsymbol{G}$ points in different directions in other grains, scattered beams emerge in each direction for which the Laue condition is satisfied in some grain by the given $\boldsymbol{G}$. Since the crystallographic orientation of the grains is assumed to have a continuous distribution in the powder sample, for a fixed $\boldsymbol{k}$ the tips of possible $\boldsymbol{k}'$s lie on the surface of a sphere of radius $|\boldsymbol{G}|$. On the other hand both the starting point and the

[12] Petrus (Peter) Josephus Wilhelmus Debye (1884-1966) was awarded the Nobel Prize in Chemistry in 1936 "for his contributions to our knowledge of molecular structure through his investigations on dipole moments and on the diffraction of X-rays and electrons in gases".

end point of the vector $\boldsymbol{G}$ have to be on the surface of the Ewald sphere determined by the vector $\boldsymbol{k}$. Therefore scattering occurs in directions that lie on the intersection line of the two spheres. This is shown in Fig. 8.15.

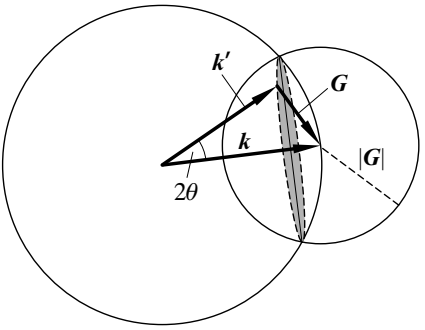


Fig. 8.15. The Ewald construction for the powder method

It is immediately seen from the figure that scattered beams emerge at angles 2θ with the incident beam for which $|\boldsymbol{G}| = 2k \sin\theta$. For such angles the scattered beams surround the direct beam conically. When recorded on a film, these appear as rings. Figure 8.16 shows the distribution of scattered radiation in a one-dimensional section.

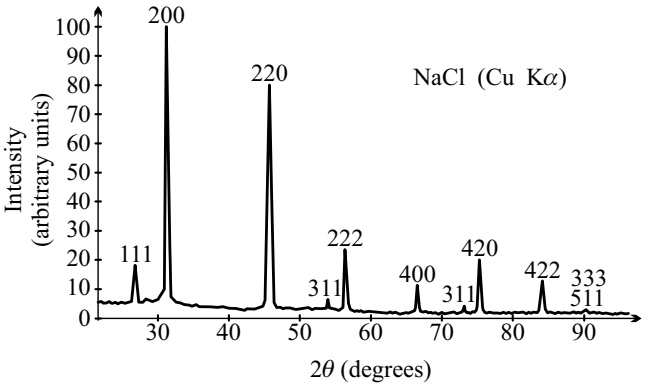


Fig. 8.16. Diffraction pattern of sodium chloride crystal, obtained with the Debye–Scherrer method. Adapted from B. E. Warren, *X-Ray Diffraction* (1969)

By measuring the angle, one can determine the length of the reciprocal-lattice vectors. This is not enough to identify a completely unknown structure – however, when the symmetries of the crystal are known, it provides one of the quickest methods for measuring the lattice constants.

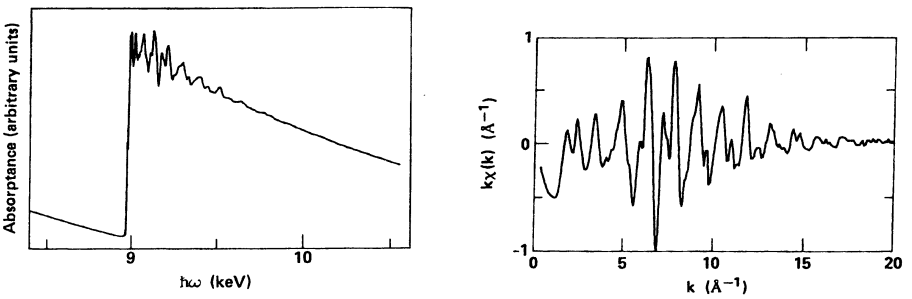
8.3 Other Methods of Structure Determination

The aim of diffraction techniques based on elastic scattering is to determine the internal structure of bulk material or the surface structure. In both cases, periodicity is assumed to extend over macroscopic distances. Besides, there exist other methods of structural analysis for the determination of the local environment of an atom. These are particularly important when the system to be examined is not regular – but they may, of course, be used to study crystalline samples, too. Probably the most adapted technique is *EXAFS spectroscopy.*[13] As its name reveals, it uses X-rays – however, they are not reflected from the sample but the fine structure of their absorption spectrum is studied.

When a photon of the incident X-ray beam is absorbed, its energy is transferred to the system of electrons. The linear absorption coefficient μ is defined through the relation

$$I = I_0 \exp(-\mu/L)\,, \tag{8.3.1}$$

where L is the thickness of the sample, while I_0 and I are the incident and transmitted intensities. In general, μ is a smooth function of the photon energy – except for the values that correspond to the ionization energies of bound states. At these energies sharp thresholds, *absorption edges* appear, since new channels become available for absorption. As shown in Fig. 8.17, the energy dependence of absorption is not really smooth in the vicinity of the absorption edge, either. EXAFS is concerned with the study of this fine structure.



(*a*) (*b*)

Fig. 8.17. (*a*) Absorption edge on the K shell of copper. (*b*) Fine structure of the absorption edge [T. M. Hayes and J. B. Boyce, *Solid State Physics*, Vol. **37**, 173 (1982)]

1–10 Å (1–10 keV) photons are energetic enough to knock out electrons even from deep levels (e.g., the 1s level). For energies in excess of the threshold value, the excited electron propagates as a spherical wave and is scattered

[13] EXAFS stands for *extended X-ray absorption fine structure.*

by adjacent atoms. These scattered waves interfere with the original wave. Depending on the photon wavelength (energy) and the geometry of the neighbors, this interference can be constructive or destructive. The energy dependence of absorption may be written as

$$\sigma(E) = \sigma_0(E)[1 + \chi(E)]\,, \tag{8.3.2}$$

where $\sigma_0(E)$ is a smooth absorption curve, and $\chi(E)$ accounts for the interference due to neighbors. Expressed in terms of the wave number rather than the energy, $\chi(E)$ is replaced by $\chi(k)$, which can be related to the Fourier transform of the radial distribution function. From measured data conclusions can be drawn about the local environment.

Like X-ray and neutron diffraction, EXAFS spectroscopy probes the interior of the sample. As it has been mentioned, electron diffraction is well suited to determining the structure of the surface layers only, since electrons do not penetrate sufficiently deeply into the sample. Nonetheless the study of solid surfaces is a fascinating subject in itself, and so several novel methods have been devised. For example, G. Binnig and H. Rohrer[14] built the scanning tunneling microscope (STM) in 1981.

The basic idea is to place the tip of a very thin stylus close to a conducting surface, at a distance of order 10 Å. The stylus itself is metallic and the tip is formed by a single atom. When a voltage is applied between the tip and the surface, electrons can tunnel through the junction, giving a measurable tunneling current. The scanning process consists of moving the tip laterally along the surface; its vertical position is controlled using piezoelectric materials. In the usual mode of operation current is kept fixed during scanning and the perpendicular deflection required to keep the current constant is monitored. The exponential dependence of the tunneling current on distance allows one to take atomic resolution images of the surface.

Obviously, the method is appropriate for the study of metallic surfaces only. For nonmetallic surfaces the atomic force microscope (AFM)[15] can be used – an instrument that measures the force between the tip of the stylus and the sample. Figure 8.18 shows the image of a surface taken by an AFM. It clearly shows that surface reconstruction gives rise to a layer with new symmetry.

[14] Gerd Binnig (1947–) and Heinrich Rohrer (1933–) shared the Nobel Prize in 1986 "for their design of the scanning tunneling microscope".

[15] G. Binnig, C. F. Quate, and Ch. Gerber, 1986.

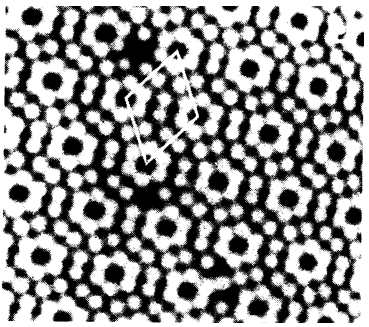


Fig. 8.18. Atomic resolution image of the (111) surface in silicon [F. Besenbacher and K. Mortensen, *Europhys. News* **21**, 68 (1990)]

Further Reading

1. J. Als-Nielsen and D. McMorrow, *Elements of Modern X-Ray Physics*, John Wiley & Sons, Ltd., New York (2001).
2. R. W. James, *The Optical Principles of the Diffraction of X-Rays*, Ox Bow Press, Woodbridge, Connecticut (1982).
3. D. McKie and C. McKie, *Essentials of Crystallography*, Blackwell Scientific Publications, Oxford (1986).
4. B. K. Vainshtein, *Modern Crystallography*, Vol. 1. *Fundamentals of Crystals. Symmetry and Methods of Structural Crystallography*, Second enlarged edition, Corrected 2nd printing, Springer-Verlag, Berlin (1996).
5. B. E. Warren, *X-Ray Diffraction*, Addison-Wesley Publishing Co., Reading, Mass. (1969).
6. M. M. Woolfson, *An Introduction to X-Ray Crystallography*, Second Edition, Cambridge University Press, Cambridge (1997).

9

The Structure of Real Crystals

In the previous chapters, where the symmetries of the crystalline state were listed and diffraction was studied, crystals were assumed to be ideal, with a strict periodicity. But real crystals are never ideal. No matter how carefully they are grown, there are always some departures from perfect order. The specific types of defects that appear in the sample and their numbers are determined, besides preparation conditions, by the history of the sample. Every single property of the material is affected by these defects – but not to the same extent. Mechanical and elastic properties of solids are most sensitive to structure. These properties cannot even be interpreted without a proper understanding of structural defects. This is why it is of the utmost importance in materials science to know the real structure of crystals, whether they occur naturally or are grown for a specific purpose.

Other properties, for example electric and magnetic properties in general, are relatively insensitive to structure although they also depend to a small degree on the presence of defects. Any deviation from the regular structure modifies the state of electrons, too, so when these are studied defects in imperfect crystals cannot be altogether ignored. In fact effects of impurities and the ensuing disorder on electronic states has become a hot research topic in the past decades leading to a lot of new discoveries. In this chapter we shall present the most characteristic types of deviation from the ideal lattice structure. The effects of defects on electronic states will be discussed in Chapters 17 and 36.

Departure from the perfectly periodic arrangement appears in crystals for various reasons and in various forms. In materials of stoichiometric composition – even when each atom is at its proper place in the primitive cell – natural isotopes are randomly distributed at the available positions, unless special preparation processes are employed. This hardly influences the properties of solids. However, when it comes to potential scattering of neutrons by nuclei, different isotopes may have highly disparate scattering lengths, and so besides the Bragg peaks an additional, smeared-out *incoherent background*

is observed because of isotopic disorder. This has to be taken invariably into account in the evaluation of measured data.

It is impossible to prepare a sample that is chemically 100% pure. Even the most effective purification techniques fail to eliminate a small amount of impurities – perhaps as little as a few parts per million (ppm). When growing crystals, impurities inevitably find their way into the sample – but they can also be introduced on purpose, to modify some property of the material that is sensitive to the presence of defects. In this case they are called *dopants* rather than impurities.

An impurity that occupies the place of an atom in the crystal is known as a *substitutional impurity*. If the size of the impurity atom is appreciably different from the original one, the lattice is locally distorted, as illustrated in Fig. 9.1. This kind of lattice distortion occurs for all other defect types, too, although this will not always be shown in the figures.

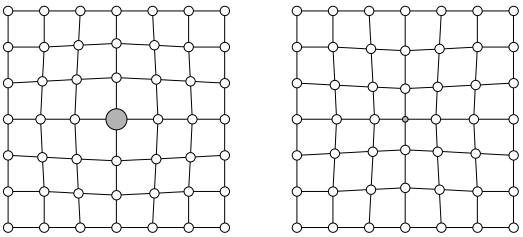


Fig. 9.1. Deformation of the lattice around a substitutional impurity atom that is larger or smaller than the original one

From a structural point of view, lattice defects that arise from the irregular arrangement of atoms are more interesting than substitutional impurities. These defects are customarily classified by the region of the sample over which they extend.

1. *Point defects* are not point-like deviations from perfect crystalline structure in the mathematical sense, however such defects do not extend over more than a few lattice constants in each direction.
2. For *line defects* deviation from the ideal structure extends over small distances along two directions, however in the third direction – along a straight or curved line – it may extend over the whole crystal.
3. For *planar defects* deviation from the ideal structure extends over a plane or a curved surface while it is limited to a few lattice constants in the third dimension. They are also called *interfacial defects* because the defect region forms an interface between two regular regions.
4. For *volume* or *bulk defects* the structure in a macroscopic three-dimensional region is different from the rest of the sample.

Below we shall discuss these four types of defect in separate sections. We shall deal with structural questions only.

9.1 Point Defects

Starting with the structure of an ideal crystal, three types of point defect are possible. One of them, an impurity atom occupying the position of an atom in the crystal has already been mentioned. Below we shall discuss the two other cases, the internal defects of pure samples.

Compared to the regular structure atoms may be missing. When a site in the lattice is vacant, one may say that a *vacancy* is present in the lattice. This situation is illustrated in Fig. 9.2(a); deformation of the lattice around the vacancy is not shown. Several references call this a *Schottky defect*, however we shall adhere to commoner usage and reserve this term for the case when two oppositely charged ions leave their sites in an ionic crystal, creating a pair of vacancies. Another type of defect is shown in Fig. 9.2(b): here an atom appears at an interstitial site, that is in the empty region among the atoms located at the sites of a regular lattice.

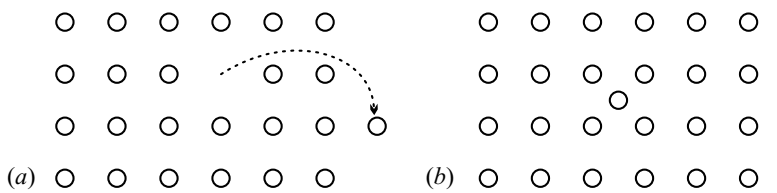


Fig. 9.2. Point defects in ideal crystal structures: (a) vacancy; (b) interstitial

If the ideal crystal is considered as the ground state of the solid, crystals with vacancies and/or interstitials can be considered as excited states of crystalline matter, since – as we shall see – in thermal equilibrium at finite temperatures these defects are present in finite concentrations.

9.1.1 Vacancies

First we shall demonstrate that vacancies are indeed generated thermally in any crystal. Consider an ideal crystal with N atoms, and assume that n vacancies are formed via the diffusion of n atoms to the surface. We shall denote the energy required to remove one atom – the formation energy of a vacancy – by ε_0; when vacancies are sufficiently far apart and thus interactions among them can be neglected, the formation energy of n vacancies is $E = n\varepsilon_0$. The total internal energy of the system is

$$E = E_0 + n\varepsilon_0\,, \tag{9.1.1}$$

where E_0 is the internal energy of the ideal crystal.

Here we are not interested in the determination of the energy ε_0. Note however that since the formation of vacancies gives rise to a change in volume, the quantity directly accessible to measurements is formation enthalpy. Vacancy formation energies for some simple metals are listed in Table 9.1. It should be noted that these are much smaller than binding energies – precisely because of the deformation of the lattice.

Table 9.1. Vacancy formation energies (in eV) and their thermal equilibrium concentrations at the melting point in some simple metals

Element	Cu	Ag	Au	Al
ε_0 (eV)	1.07	1.09	0.98	0.78
c_v	4.5×10^{-5}	3.5×10^{-5}	17×10^{-5}	6×10^{-5}

In spite of the increase in energy due to the vacant sites, the concentration of vacancies remains finite in thermal equilibrium at finite temperatures as the increase in energy is compensated for by the increase in entropy due to disorder, and thus free energy will be lower than in the ordered state. The configurational entropy of the disordered state can be determined by assuming that the atoms missing from the vacancies have diffused to the surface and occupy sites that can be considered as the continuation of the crystal. A sample containing N atoms and n vacancies can therefore be considered to have $N+n$ lattice sites, of which n are vacant. Assuming that the n vacancies are distributed randomly over the $N+n$ sites, the entropy associated with disorder is

$$S_\mathrm{config} = k_\mathrm{B} \ln \binom{N+n}{n} = k_\mathrm{B} \ln \frac{(N+n)!}{N!\,n!}\,. \tag{9.1.2}$$

When besides the number of lattice sites the number of vacancies is also large, the Stirling formula (C.3.29)

$$\ln n! = \left(n + \tfrac{1}{2}\right) \ln n - n + \tfrac{1}{2} \ln 2\pi + \mathcal{O}\,(1) \tag{9.1.3}$$

can be used. Keeping only leading-order terms,

$$S_\mathrm{config} \approx k_\mathrm{B} N \ln \frac{N+n}{N} + k_\mathrm{B} n \ln \frac{N+n}{n}\,. \tag{9.1.4}$$

In addition to this, the finite temperature entropy S_0 of the perfectly regular crystal needs to be taken into account.

The formation of vacancies is accompanied by a change in the sample volume as atoms move to its surface. Thermal equilibrium therefore occurs at

the minimum of the Gibbs free energy (sometimes also known as free enthalpy or Gibbs potential)

$$G = E - TS + pV\,. \tag{9.1.5}$$

Provided that the displacement of atomic positions about the vacancies is a small perturbation that can be ignored, the total volume of the sample increases from Nv_0 to $(N+n)v_0$, where v_0 is the specific volume per atom. Substituting the expressions for internal energy, entropy, and volume into the Gibbs free energy formula,

$$\begin{aligned} G \approx E_0 + n\varepsilon_0 - TS_0 - k_\mathrm{B}TN\ln\frac{N+n}{N} - k_\mathrm{B}Tn\ln\frac{N+n}{n} \\ + p(N+n)v_0\,. \end{aligned} \tag{9.1.6}$$

To obtain the equilibrium number of vacancies this expression has to be minimized:

$$\frac{\partial G}{\partial n} \approx \varepsilon_0 - k_\mathrm{B}T\ln\frac{N+n}{n} + pv_0 = 0\,. \tag{9.1.7}$$

If the number of vacancies is much smaller than the number of lattice sites ($n \ll N$), we have

$$n = N\mathrm{e}^{-(\varepsilon_0+pv_0)/k_\mathrm{B}T}\,. \tag{9.1.8}$$

Vacancy formation energies are on the order of an eV, so pv_0 can be neglected compared to them:

$$n = N\mathrm{e}^{-\varepsilon_0/k_\mathrm{B}T}\,. \tag{9.1.9}$$

The equilibrium number of thermally generated vacancies shows strong temperature dependence. As indicated by the data in the last line of Table 9.1, close to the melting point the concentration of such defects is on the order of 10^{-4}–10^{-5}. If the sample contained only thermally generated vacancies, their number would be entirely negligible at room temperatures because of the exponential temperature dependence. When a sample is cooled quickly from high temperatures (quenching), high nonequilibrium vacancy concentrations may freeze into it. A part of them can be removed by annealing.

When thermally excited atoms diffuse from the interior of the crystal to the surface leaving vacancies behind, the dimensions of the crystal increase more rapidly than the lattice constant (the microscopic distance between atoms at regular lattice sites). This is why different thermal expansion coefficients are obtained from diffraction measurements that are sensitive to atomic distances and variations of the macroscopic dimensions of the sample.

If the crystal is built up of more than one kind of atom, and vacancies appear on each sublattice independently, then the thermal vacancy concentration can be determined separately for each kind because formation energies depend on the atom and its environment. If removing a jth type atom requires an energy ε_j – and, as before, the contribution due to the change in volume can

be neglected – then the Gibbs free energy of a sample containing n_j vacancies of the jth type is

$$G = E_0 + \sum_j n_j \varepsilon_j - TS_0 - \sum_j k_\mathrm{B} T N_j \ln \frac{N_j + n_j}{N_j} - \sum_j k_\mathrm{B} T n_j \ln \frac{N_j + n_j}{n_j} \,, \tag{9.1.10}$$

where N_j is the number of jth type atoms in the crystal. (The same expression gives the Helmholtz free energy as the volume term is absent.) This expression is minimized by

$$n_j = N_j \mathrm{e}^{-\varepsilon_j / k_\mathrm{B} T} \,. \tag{9.1.11}$$

Obviously, defects with the lowest formation energy are the most abundant. In most cases only these need to be taken into account.

9.1.2 Interstitials

So far it has been assumed that when a vacancy is generated the atom moves to the surface of the crystal. This is often not the case: it may also occupy a position that is not occupied in a perfect crystal. We have seen that by considering atoms in a crystalline structure as spheres there may be large enough regions among them for a new atom to fit in (especially when neighboring atoms are slightly displaced). The narrow space among the original atoms is an interstice, so an atom occupying it is called an *interstitial atom* or simply an *interstitial*.

The energy of an interstitial atom depends on its exact position. Local minima of the potential due to neighboring atoms are to occur at positions that are surrounded symmetrically by atoms of the regular lattice. For a simple cubic crystal such are the body center and the face centers of the cube. In face-centered cubic crystals two such characteristic sites exist, as shown in Fig. 7.13. One of them is the center of the cube at $\frac{1}{2}\frac{1}{2}\frac{1}{2}$; the second is the center of an octant, at $\frac{1}{4}\frac{1}{4}\frac{1}{4}$ (which quadrisects the space diagonal) or any equivalent position. In relation to Fig. 7.13 we also saw that the local environments around the interstitial are different at the two sites. In the first case the interstitial atom is surrounded by six first neighbors in an octahedral geometry, thus the site has an octahedral (cubic) symmetry. The local point group of reflections and rotations that leave this point invariant is the group O_h. In the second case the defect site has four nearest neighbors that are arranged tetrahedrally. Local symmetries are elements of the tetrahedral point group T_d. As mentioned in Chapters 6 and 7, the pattern of energy level splitting may be different for atoms in the two environments. This can be used for their identification, through optical spectroscopy or nuclear magnetic resonance.

Two interstitial sites are distinguished in a body-centered cubic lattice, too. As shown in Fig. 7.8, two empty sites of high symmetry are found among the

atoms at the vertices. Edge centers (at $\frac{1}{2}00$ and equivalent positions, including face centers) are called octahedral sites – although the six neighbors do not form a regular octahedron and the symmetry is only tetragonal. They are also called "small" sites because they offer relatively little space. The lattice needs to be deformed to a lesser extent to make room for an interstitial atom when the atom occupies one of the tetrahedral sites, $\frac{1}{4}\frac{1}{2}0$ or an equivalent position. These sites are called "large".

The equilibrium concentration of interstitials can be evaluated along the same lines as that of vacancies, and the result (9.1.9) applies to this case, too. However, the theoretical determination of the formation energy is now much more difficult than for vacancies: an interstitial deforms the lattice more than a vacancy does, as it has to make some room for itself among other atoms, and so the contribution due to the deformation of the crystal is more important. Because of the larger deformation the formation energy is also somewhat higher, on the order of a few eV, consequently interstitials are more difficult to generate thermally than vacancies.

Further types of atomic configurations are possible when the additional atom is allowed to displace one or several neighbors substantially. An important case is when an atomic position is occupied by two atoms symmetrically on either side of the site in a regular crystal. This is in fact an interstitial pair, and the configuration is called a *split interstitial*. Such situations are shown in Fig. 9.3. For example, in the primitive cell of a body-centered cubic crystal the atom originally at the body center and the additional atom may be arranged symmetrically on both sides of the center, in the $\langle 110 \rangle$ direction. In a possible split interstitial configuration of a face-centered cubic lattice an atom at a face center is displaced in the $\langle 100 \rangle$ direction to make enough space for another atom on the other side of the face.

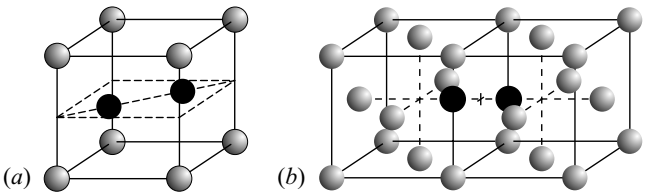


Fig. 9.3. Split interstitials: (*a*) in the [110] direction in a body-centered cubic lattice; (*b*) in the [100] direction in a face-centered cubic lattice

In copper, the prototype of face-centered cubic crystals, this defect – a split interstitial along the $\langle 100 \rangle$ direction – has a lower energy than the configuration in which a copper atom is simply placed at the interstitial site.

While in these dumbbell defects two atoms share the space of one in a perfect crystal – making, of course, some extra space for themselves among the neighbors –, it may also happen that three atoms share two lattice positions, or

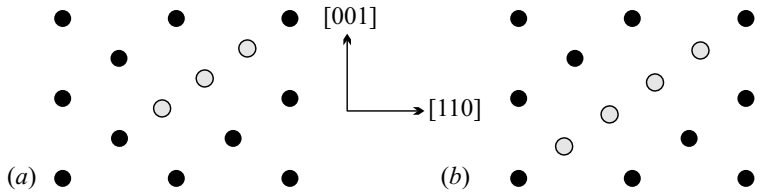


Fig. 9.4. Crowdion defects in a body-centered cubic crystal along the [111] direction: (*a*) three atoms sharing two lattice sites; (*b*) four atoms sharing three lattice sites

four atoms share three lattice positions, etc. along a line. Such configurations extending over a few atomic distances in one direction are called *crowdions*. Figure 9.4 shows the atomic arrangement in a plane section of a body-centered cubic crystal cubic crystal containing $C_{3/2}$ and $C_{4/3}$ types of crowdion along the [111] direction.

Crowdions tend to have higher energies than split interstitials, and thus occur much more rarely naturally.

9.1.3 Pairs of Point Defects

As we have seen, in thermal equilibrium the concentrations of vacancies and interstitials are both finite and small enough for treating them as if each defect had been generated independently of the others. Strictly speaking this is not true. When vacancies are thermally generated, some of the atoms that leave their lattice sites move to the surface, while others become trapped at interstitial sites. Which of these possibilities occurs is of particular interest in ionic crystals. A missing ion breaks local charge neutrality, which has to be restored by another nearby defect, since charge neutrality must be valid in any relatively small region of space otherwise the Coulomb energy would become excessively large and the configuration would be energetically highly unfavorable. Restoration of charge neutrality can occur via the formation of *Schottky defects* or *Frenkel defects*.

Schottky Defects

It was first pointed out by W. Schottky (1930) that in ionic crystals built up of positively and negatively charged ions a vacancy appearing on the sublattice of cations is accompanied by a nearby vacancy on the sublattice of anions. Although both ions diffuse to the surface, charge neutrality is maintained locally inside the sample. This defect – a pair of oppositely charged nearby vacancies – is called a Schottky defect (although several references use this term for a single vacancy). The alternative terms *vacancy pair* and *divacancy* are also widely used. This type of defect is illustrated in Fig. 9.5.

Equation (9.1.9), or its generalization, (9.1.11) cannot be used to determine the concentration of Schottky defects as the derivation of these formulas was

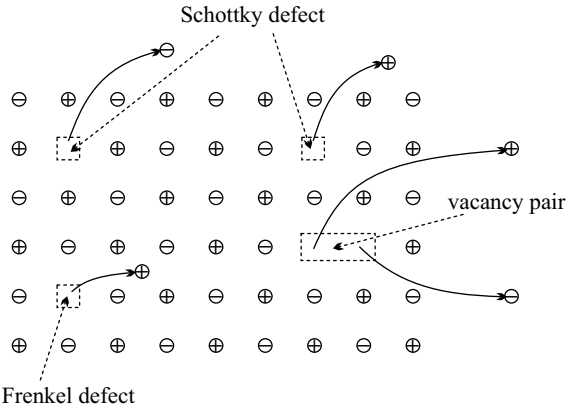


Fig. 9.5. Schottky and Frenkel defects in an ionic crystal

based on the independence of individual defects. When the charge of the jth type ion is q_j, charge neutrality requires

$$\sum_j q_j n_j = 0\,. \tag{9.1.12}$$

When minimizing the Gibbs free energy (9.1.10), this can be taken into account by a Lagrange multiplier. The quantity to be minimized is then

$$G + \lambda \sum_j q_j n_j\,. \tag{9.1.13}$$

Along the same lines as above, the calculation yields

$$n_j = N_j \mathrm{e}^{-(\varepsilon_j + \lambda q_j)/k_\mathrm{B}T}. \tag{9.1.14}$$

The so-far undetermined multiplier λ has to be determined from the equation for charge neutrality,

$$\sum_j q_j N_j \mathrm{e}^{-(\varepsilon_j + \lambda q_j)/k_\mathrm{B}T} = 0\,. \tag{9.1.15}$$

For simplicity, consider a crystal built up of an equal number N of oppositely charged anions (q) and cations ($-q$). When vacancy formation energies are denoted by ε_+ and ε_-, (9.1.14) tells us that the equilibrium numbers n_+ and n_- are given by

$$n_+ = N\mathrm{e}^{-(\varepsilon_+ + \lambda q)/k_\mathrm{B}T}, \qquad n_- = N\mathrm{e}^{-(\varepsilon_- - \lambda q)/k_\mathrm{B}T}. \tag{9.1.16}$$

The condition for neutrality is now $n_+ = n_-$, which implies $\lambda q = (\varepsilon_- - \varepsilon_+)/2$, and so

$$n_+ = n_- = N\mathrm{e}^{-(\varepsilon_+ + \varepsilon_-)/2k_\mathrm{B}T}. \tag{9.1.17}$$

The number of Schottky defects is thus determined by the average of the two formation energies. As listed in Table 9.2, formation energies are typically around 2 eV for alkali halides, i.e., much smaller than the binding energy of the ion pair, on the order of 10 eV.

Table 9.2. Formation energies of Schottky and Frenkel defects (in eV) in some simple ionic crystals

Schottky defects		Frenkel defects	
Compound	ε_0 (eV)	Compound	ε_0 (eV)
LiF	2.51	AgCl	1.4
NaCl	2.28	AgBr	1.1
NaBr	1.72	CaF_2	2.6
KCl	2.28	SrF_2	2.3
KI	1.60	BaF_2	1.9
CsCl	1.86	$SrCl_2$	1.9

Since we shall not discuss in detail the properties of ionic crystals elsewhere, it should be mentioned here that in those ionic crystals (e.g., alkali halides), where Schottky defects are the most important defects, electric conductivity can be interpreted in terms of the motion of vacancies. As massive ions move slowly, resistivity is several orders of magnitude larger than in ordinary metals. Compared to metals, where resistivity is typically on the order of $\mu\Omega\,\mathrm{cm}$, values between 10^2 and $10^8\,\Omega\,\mathrm{cm}$ are observed in ionic crystals. Because of the thermal generation of vacancies electrical conductivity increases with temperature, showing thermally activated behavior.

The fact that vacancies and interstitials carry charge plays an important role in determining the properties of ionic crystals. If a negatively charged ion is missing, the hole has an effective positive charge, and can therefore bind an electron, restoring charge neutrality. The defect consisting of an electron bound to a vacancy is called a *color center* or *F-center*.[1] Defects in which two or three electrons are bound by two or three neighboring vacancies are called *M*- and *R-centers*.[2]

[1] The name refers to the property that electrons in the attractive potential of the vacancy are on hydrogen-like orbitals, so they can absorb light of specific frequencies, giving a particular color to the crystal. The name *F*-center comes from the German word for color, *Farbe*.

[2] *M*-centers are also called di-*F*-centers or F_2-centers.

Frenkel Defects

When a vacancy is generated, the atom may not diffuse to the surface but become trapped at a nearby interstitial site, as shown in Fig. 9.5. In ionic crystals the complex of a vacancy and a nearby interstitial of the same ion (which thus carry opposite effective charges) is called a *Frenkel defect.*[3]

The equilibrium concentration of Frenkel defects can be determined along the same lines as the concentration of simple vacancies. Suppose the sample has N atomic positions, of which n are left vacant. Since the atoms do not diffuse to the surface but become trapped at interstitial sites, the volume does not change in a first approximation. We shall denote the number of possible interstitial sites by N_{int}. If the n vacancies are formed randomly at the N atomic positions and the displaced ions occupy the interstitial sites at random as well, the configurational entropy is given by

$$S_{\text{config}} = k_{\text{B}} \ln \left[\frac{N!}{(N-n)!n!} \frac{N_{\text{int}}!}{(N_{\text{int}}-n)!n!} \right]. \tag{9.1.18}$$

If the formation energy of a Frenkel defect is ε_0, minimization of the free energy leads to

$$n = \sqrt{N N_{\text{int}}}\, \exp(-\varepsilon_0 / 2k_{\text{B}}T) \tag{9.1.19}$$

in the $N, N_{\text{int}} \gg n$ limit.

Unlike alkali halides, silver halides contain such defects. After the formation of vacancies on the sublattice of silver ions, these positively charged ions become trapped at interstitial sites. Frenkel defects are also found in alkaline-earth fluorides. Formation energies are listed in Table 9.2.

In principle, another type of defect pair can exist, in which all ions sit on lattice sites but a cation and a nearby anion swap positions. In ionic crystals this is highly prohibited by the unfavorable effective charge that would appear and the large energy involved. However, such defects are indeed observed in intermetallic compounds, where swapping between different components entails a smaller increase in energy. They are called *antisite defects.* More generally, an antisite defect A_{B} is an atom of type A that occupies a wrong site on sublattice B. Such individual antisite defects are of special interest in doped semiconductor compounds.

9.2 Line Defects, Dislocations

Atoms in a crystal may become misplaced with respect to their regular position along a whole line because of very large internal mechanical strains and stresses occurring while working the material. When a crystal is bent or twisted, atomic layers may slip on each other in such a way that the majority of the atoms around the slip – except for the vicinity of a line – find

[3] J. Frenkel, 1926.

themselves in another locally aligned environment. Defects of this kind may develop during crystal growth as well. They are called *dislocations*.

As a large number of bonds are broken between atoms along the line, the energy of a dislocation is much higher than that of a vacancy or an interstitial, consequently dislocations are not generated thermally, only during mechanical working. The density of dislocations depends strongly on the preparation conditions. In very carefully grown germanium or silicon crystals they may be as low as 10^2–$10^3/\mathrm{cm}^2$, while in metals subject to high mechanical strains they may range up to 10^{10}–$10^{12}/\mathrm{cm}^2$. Their presence strongly affects the mechanical properties of materials; they also play an important role in the growth of crystals. However, we do not discuss these points any further here, as we focus on the issues of structure only.

9.2.1 Edge and Screw Dislocations

The structure around line defects is most easily illustrated using the Volterra construction (or Volterra process).[4] An imaginary cut is made into a crystal, terminating in its interior on a line called the *dislocation line*. Its shape being irrelevant, the cut is chosen, for simplicity, to be in a flat plane. In the next step the two parts separated by the cut are displaced relative to one another, allowing elastic deformations of the lattice near the dislocation line. If the relative displacement is a translation vector of the lattice, then – except for the region close to the dislocation line – the two parts are in registry, and can be joined back together to restore the crystalline order. Naturally, if the displacement vector is not in the cut plane, matter has to be added to or removed from the crystal. The Volterra construction is illustrated schematically in Figs. 9.6 and 9.7 showing the crystal before and after the displacement for two different cases.

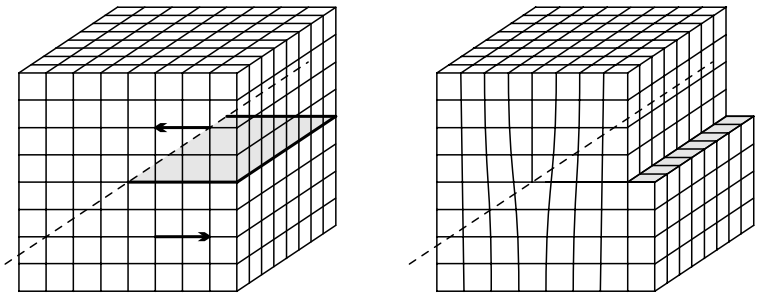


Fig. 9.6. Volterra construction of an edge dislocation: a part of the crystal is slipped along a cut. The arrow shows the direction of slip, which is perpendicular to the dislocation line

[4] V. Volterra, 1907.

Figure 9.6 shows the case where the displacement is perpendicular to the dislocation line. It is clearly seen that the defect – the "imperfect" region – is limited to the vicinity of the dislocation line. This type of line defect is called an *edge dislocation*.[5]

When the crystal is considered as an elastic medium, one side is compressed and the other is stretched, giving rise to elastic stress. The energy of dislocation can be determined from this elastic energy.

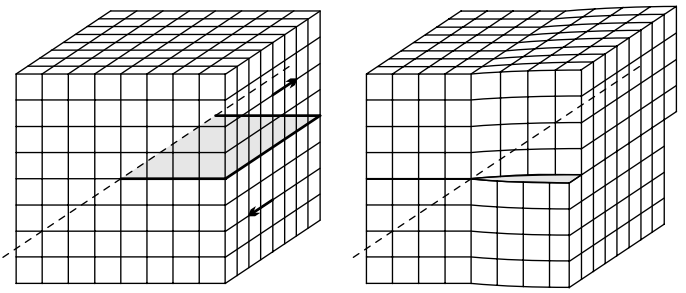


Fig. 9.7. Formation of a screw dislocation: an atomic layer is slipped on another in a part of the sample

The situation is slightly different when the two parts are slipped relative to each other parallel to the dislocation line, as shown in Fig. 9.7. The arrangement of atoms in the layers below and above the cut plane is shown in Fig. 9.8(a). Misalignment is again limited to the vicinity of a line. It is clear from the figure that by making a full turn around the dislocation line in a crystal plane perpendicular to it, one arrives at the atomic layer just below or just above. If traveling around is continued in the same way a spiral, a screw is traced out. This type of defect is therefore called a *screw dislocation* or sometimes a *Burgers dislocation*.[6]

More general dislocations may arise when the dislocation line is not straight. In such cases it will be parallel or perpendicular to the slip vector only exceptionally. An example is shown in Fig. 9.8(b), with the atomic arrangement in the layers below and above the cut plane. The atomic arrangement shown at the bottom of the figure is similar to that of a screw dislocation, however it is transformed into an edge dislocation inside the sample. This type of dislocation is called a mixed dislocation.

[5] Edge dislocations are sometimes called *Taylor–Orowan dislocations*, although the concept of dislocation was introduced in 1934 by three researchers independently of each other, E. Orowan, M. Polányi, and G. I. Taylor.

[6] J. M. Burgers 1939.

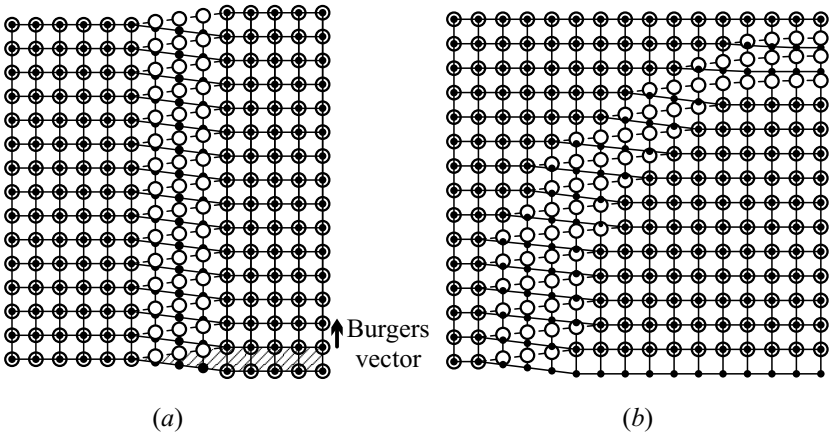


Fig. 9.8. Atomic positions above (large empty circles) and below (small full circles) the cut plane for (*a*) a screw dislocation; (*b*) a screw dislocation being transformed into an edge dislocation (mixed dislocation)

9.2.2 The Burgers Vector

Dislocations are characterized first by the dislocation line – along which regular order is disrupted –, and second by the vector specifying the relative displacement of the parts below and above the cut. However, as it has been mentioned, the cut plane does not have any physical significance. The Volterra construction of line defects served only as an illustration of dislocations. To unambiguously specify the displacement of atomic layers we adopt the method introduced by J. M. Burgers in 1939. First, a direction is associated with the dislocation line and a lattice point S that is distant from the dislocation is selected. Then, starting from S and staying all the time in the region where the local atomic environment is undistorted, a closed path encircling the dislocation is traversed, in counterclockwise sense when viewed from the chosen direction. The path may be chosen at will, it should just contain the same number of steps to the left as to the right, upward as downward, and forward as backward – more precisely, the number of steps in the direction of each primitive vector $\boldsymbol{a}_i$ $(i = 1, 2, 3)$ should be the same as in the opposite direction $-\boldsymbol{a}_i$. The traversed path is called the *Burgers circuit*. In crystals free of dislocations any such path terminates on the starting point S. However, when a dislocation is encircled, the path fails to close, as shown in Fig. 9.9 in the plane perpendicular to the dislocation line. No matter how the Burgers circuit is chosen for a fixed starting point S, it always terminates on the same lattice point F, which is now different from S. Displacement through a further – not necessarily primitive – lattice vector is required to close the circuit around the dislocation, and thus get back from F to S. The lattice vector pointing from F into S is called the *Burgers vector* of the dislocation.

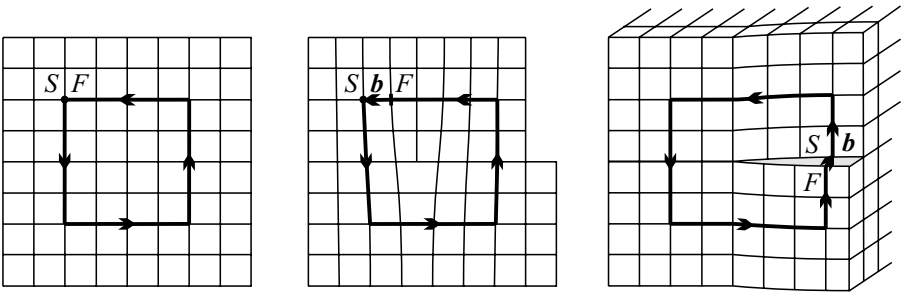


Fig. 9.9. The Burgers circuit and the Burgers vector (a) in a regular crystal; (b) around an edge dislocation; (c) around a screw dislocation

As shown in Fig. 9.9(b), the Burgers vector is perpendicular to the direction of the dislocation line for edge dislocations, while for screw dislocations one or more extra steps have to be made along the dislocation line to have the Burgers circuit closed. The Burgers vector is therefore parallel to the dislocation line in this case.

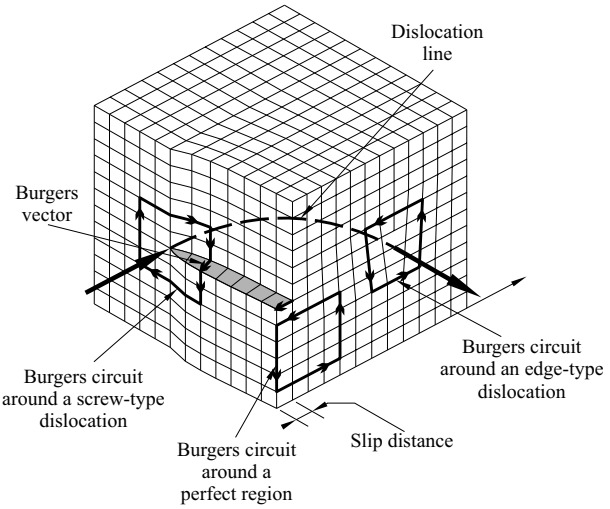

Dislocation line
Burgers vector
Burgers circuit around an edge-type dislocation
Burgers circuit around a screw-type dislocation
Slip distance
Burgers circuit around a perfect region

Fig. 9.10. Three Burgers circuits in a crystal with mixed dislocation: in a defect-free region, around the dislocation line in a region where it is of screw-type, and where it is of edge-type

We have seen that the dislocation line is not always straight: it may change its direction in the crystal – and along with it its character, too. Even so, regardless of the particular choice of the Burgers circuit, as long as it encircles

the dislocation line, the same Burgers vector is obtained. This is illustrated in Fig. 9.10. The dislocation shown here looks like a screw dislocation from the front and an edge dislocation from the right – nevertheless regardless of the particular choice of the Burgers circuit, the same Burgers vector is obtained when the dislocation line is encircled in the same sense.

9.2.3 Dislocations as Topological Defects

The independence of the Burgers vector of the circuit encircling the dislocation indicates that it is a topological property of the atomic arrangement around the dislocation. For a precise illustration consider a two-dimensional section of a crystal with a dislocation and a Burgers circuit Γ in it. Atomic positions in the perfect crystal are marked next to the actual atomic positions in Fig. 9.11.

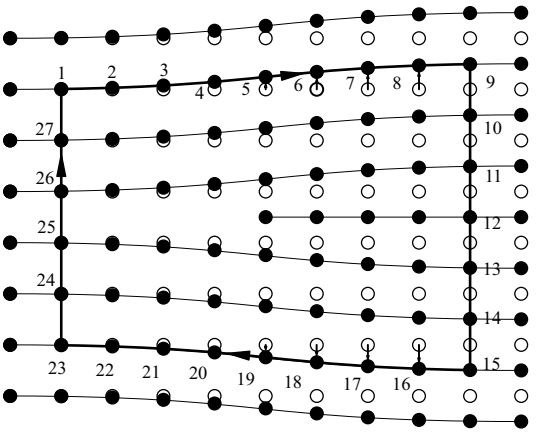


Fig. 9.11. Displacement of the atoms on a Burgers circuit in a plane perpendicular to the dislocation with respect to their positions in a perfect crystal

Along the Burgers circuit, where the lattice shows an almost perfect periodicity, the displacement caused by the dislocation can be defined as the vector from the closest lattice point of the ideal crystal to the actual atomic position. Thus by definition all possible displacements lie inside the Wigner–Seitz cell. These displacements are shown in Fig. 9.12(a). As the Burgers circuit is traversed, the displacement of subsequent lattice points increases; it jumps to an equivalent value (from $u_x = a/2$ to $u_x = -a/2$) when the boundary of the Wigner–Seitz cell is reached, and then approaches zero. The sequence of points in the (u_x, u_y) plane can be viewed as a mapping of the Burgers circuit (defined in real space) to displacement space.

Far from the dislocation core – i.e., the dislocation line – the displacement varies only slightly from one lattice point to the next, and so a continuous displacement field $\boldsymbol{u}(\boldsymbol{r})$ may be defined. While traversing the Burgers circuit

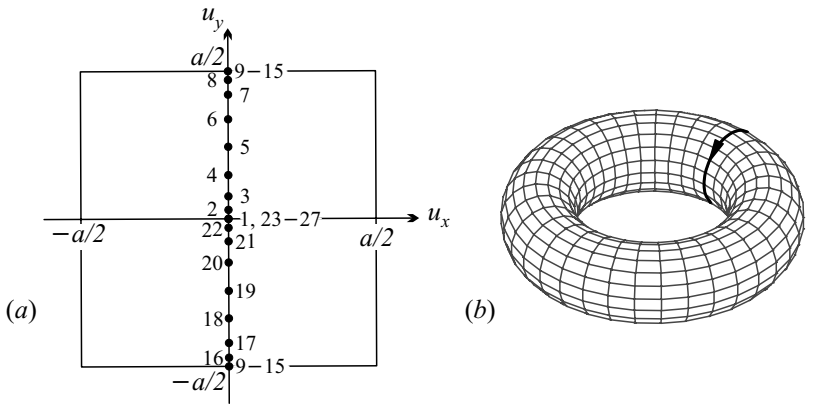


Fig. 9.12. (*a*) Displacement of the atoms around the Burgers circuit in Fig. 9.11 in the Wigner–Seitz cell of the lattice. (*b*) Mapping of the Burgers circuit into the torus in displacement space with periodic boundary conditions

around the dislocation in real space, the displacements $\boldsymbol{u}(\boldsymbol{r})$ trace out a piecewise continuous closed path in displacement space. The discontinuities arise from jumps between equivalent boundary points of the Wigner–Seitz cell.

When the derivative of the displacement field with respect to the arc length of the Burgers circuit is integrated along the circuit, the result is nonzero even though the integration contour is a closed path.[7] The reason for this is that jumps introduced by the imposed periodic boundary conditions are not included in the integrand. The integral is the Burgers vector of the enclosed dislocation,

$$\oint_{\Gamma} \frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}s}\,\mathrm{d}s = \boldsymbol{b}\,, \tag{9.2.1}$$

independently of the particular choice of the circuit Γ. The topological character of this contour integral is better illustrated when the displacement field is defined on a torus, where periodic boundary conditions are automatically taken into account. In the example discussed above the curve obtained by mapping the Burgers circuit into displacement field goes around the torus once, as shown in Fig. 9.12(*b*).

When positions of the atoms lying along the Burgers circuit are continuously modified, or when the Burgers circuit is chosen differently, deforming it in small steps, the image of the Burgers circuit in displacement space is also continuously deformed. However, when drawn on a torus, the topological property of going around it once cannot change. This means that no sequence of small displacements can make the image curve shrink to a point – which

[7] In the continuum limit it does not matter whether the integration path is the open Burgers circuit or a closed contour obtained by closing the Burgers circuit with the Burgers vector.

is the image of a Burgers circuit in a perfect crystal. When the Burgers vector is twice as long, the path in displacement space goes around the torus twice. When the Burgers vector is perpendicular to the one in the figure, the path encircles the hole of the torus; this contour cannot be shrunk to a point either by small distortions of atomic positions. Determining the Burgers vector is equivalent to specifying how many times the path goes around the torus in each direction. These winding numbers can be considered to be the topological quantum numbers of the dislocation. For three-dimensional crystals displacement is a three-dimensional vector that is defined in the crystal's Wigner–Seitz cell, or, to visualize the topology better – by exploiting the lattice-periodic character of displacement space – on a 3D torus. Dislocations can therefore be characterized by three winding numbers that correspond to the three components of the Burgers vector.

Since dislocations are characterized by a topologically determined Burgers vector, open dislocation lines cannot exist inside the sample. They either reach the surface of the sample or close back on themselves, forming dislocation loops. As a matter of fact, an even stronger statement can be made. When a dislocation with a Burgers vector $\boldsymbol{b}$ splits into two dislocations with Burgers vectors $\boldsymbol{b}_1$ and $\boldsymbol{b}_2$, the following conservation theorem applies to them:

$$\boldsymbol{b} = \boldsymbol{b}_1 + \boldsymbol{b}_2 \,. \tag{9.2.2}$$

At high dislocation densities interactions between dislocations may become important, giving rise to dislocation splitting. Nevertheless the previous relationship of Burgers vectors must be valid at each splitting.

9.2.4 Disclinations

When traversing a closed path that encircles a dislocation, everywhere except for the immediate vicinity of the dislocation the local environment of each atom is found to be identical to that in a regular lattice, moreover the edges of the primitive cells are all parallel. There exist other types of defects for which this is not true. One such possibility may be visualized by making two cuts at an angle φ in the Volterra construction, removing the part of the crystal between them, and then joining the two cut surfaces. Another may be conceived as the result of cutting the crystal in one place, folding the two cut surfaces an angle φ apart, and then filling the empty region with aligned atomic layers. Crystal structures featuring such defects are shown in parts (*a*) and (*b*) of Fig. 9.13. Though the defect is once again limited to a line, it would be reflected in the shape of the entire crystal. Because of the orientational misalignment F. C. Frank (1958) called this defect a *disclination*.[8]

As a wedge-shaped region has been cut out of or inserted into a regular crystal, this type of defect is called a *wedge disclination*. If the region around

[8] From the Greek word $\kappa\lambda\iota\nu\omega$ (klino) = lean, bend.

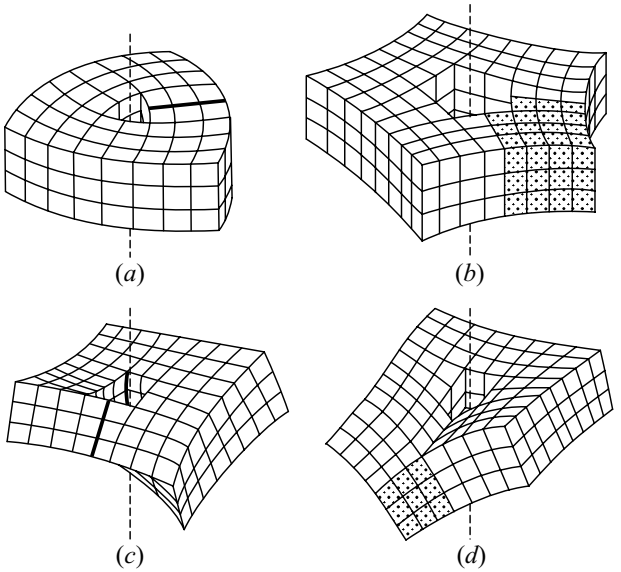


Fig. 9.13. Theoretically possible disclinations in crystals: (*a*) and (*b*) wedge disclinations, (*c*) and (*d*) screw disclinations

the disclination core were also shown, the reader would see that the coordination number there is one more or one less than elsewhere. Parts (*c*) and (*d*) show screw disclinations. In this case the finite angular range in which atoms are removed or inserted is not parallel but perpendicular to the disclination line.

Such defects are of academic interest in crystalline solids. However, a wide variety of disclinations may appear in liquid crystals because of their smaller mechanical rigidity and the presence of crystalline order along one or two directions only. An exhaustive listing and rigorous mathematical discussion of all possible types of disclination requires the apparatus of algebraic topology. Figure 9.14 shows the arrangements of columns around longitudinal wedge disclinations of angles $-\pi/3$ and $\pi/3$ in the discotic columnar phase. The

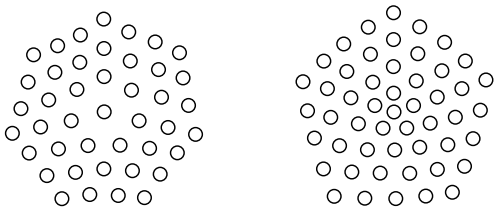


Fig. 9.14. Wedge disclinations in a columnar liquid crystal. Columns are shown in a top view

central column is surrounded by five others in the first and seven others in the second case.

9.2.5 Dislocations in Hexagonal Lattices

A special situation arises in hexagonal structures – as well as in face-centered cubic ones, where atoms in the (111) plane are arranged hexagonally. Figure 9.15 shows yet another Volterra construction in a two-dimensional hexagonal crystal. The crystal is cut along two lines that are 60° apart and have a common end point (in three dimensions: dislocation line), and then all atoms inside the wedge-shaped region are displaced in the direction perpendicular to $\boldsymbol{a}_1$ so that each atom moves exactly one line forward.

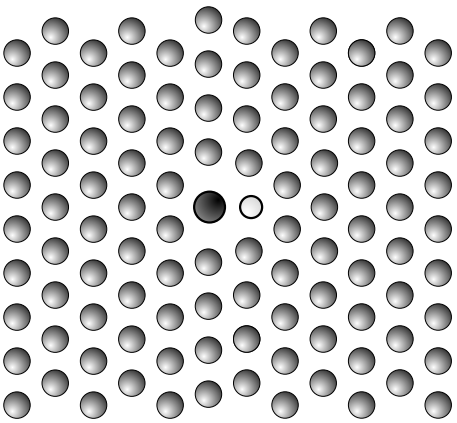


Fig. 9.15. Atomic arrangement around the dislocation in a hexagonal structure. The coordination number is seven and five for atoms marked by large and small circles, respectively

As it can be seen, the Burgers circuit can be closed by the vector $\boldsymbol{a}_1$, so this is the Burgers vector of the dislocation. The same atomic arrangement arises when the crystal is cut halfway along the direction of the primitive vector $\boldsymbol{a}_2$ and a new line of atoms is inserted. Since $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ are at 120° to each other, the atoms are arranged in such a way as if in addition to the line of atoms along the $\boldsymbol{a}_2$ direction another one, along $\boldsymbol{a}_1 + \boldsymbol{a}_2$ had also been inserted. When one takes a closer look at the region around the dislocation core, the atom at the tip of the wedge is seen to have five nearest neighbors, while the one next to it has seven. This dislocation is thus the same as two nearby disclinations, one at −60° and the other at 60°. As we shall see in the next section, dislocations of a face-centered cubic crystal can also be considered as complex defects.

9.3 Planar Defects

It happens frequently during crystal growth that instead of being stacked perfectly, subsequent atomic layers are somewhat misplaced, giving rise to a defect over an extended two-dimensional surface. The same can easily occur as a result of mechanical strain. Such defects are called *stacking faults*. If the stacking fault does not extend to the surface of the sample new kinds of line defects may appear where it terminates.

Misalignment does not necessarily occur in a plane. When crystal growth starts independently in several nuclei, the orientation of the axes around the nuclei will be random, i.e., independent of other nuclei. This is how polycrystalline materials are formed. Long-range order is limited to finite-sized macroscopic domains called grains or crystallites. Misalignment occurs only at two-dimensional grain boundaries. Such two-dimensional defects are called *grain boundaries*.

9.3.1 Stacking Faults

Structures with hexagonal symmetry were presented in Section 7.3. We saw that depending on the way layers with sixfold symmetry are stacked, various structures (simple, close-packed, double close-packed) may arise. In a hexagonal close-packed (hcp) structure the stacking order along the sixfold axis is $ABABAB\ldots$, while in a double close-packed structure it is $ABACABAC\ldots$. It may happen that a plane does not get into its proper position. For example, from a certain point onwards C-type layers may appear instead of B-type layers; in this case the previous stacking order $ABABAB\ldots$ of the hcp structure is replaced by $ACACAC\ldots$. It may equally happen that the stacking order $ABABAB\ldots$ suddenly switches to $CBCBCB\ldots$.

We have seen that the (111) planes in a face-centered cubic lattice also possess hexagonal symmetry. A part of such a plane with some atomic coordinates is shown in Fig. 9.16.

If the atomic layer shown in the figure is called A-type, its translation through $\frac{1}{6}[11\bar{2}]$, $\frac{1}{6}[\bar{2}11]$, or $\frac{1}{6}[1\bar{2}1]$ gives a B-type configuration. When the same translation is performed on a B-type layer, a C-type layer is obtained. In an fcc structure the stacking order along the [111] direction is $ABCABCABC\ldots$. Figure 9.17(a) shows the stacking of close-packed layers in another sectional view. The crystal is cut perpendicular to the direction $[1\bar{1}0]$; atoms coplanar with the origin are marked by empty circles, while atoms in the neighboring planes by full circles.

When a face-centered cubic crystal is deformed by slipping the part of the crystal above an A-type layer through any of the three vectors given in the previous paragraph, a stacking fault arises, since a B-type layer is now missing, and the new sequence is $\ldots ABCACABC\ldots$. The atomic arrangement is shown in Fig. 9.17(b).

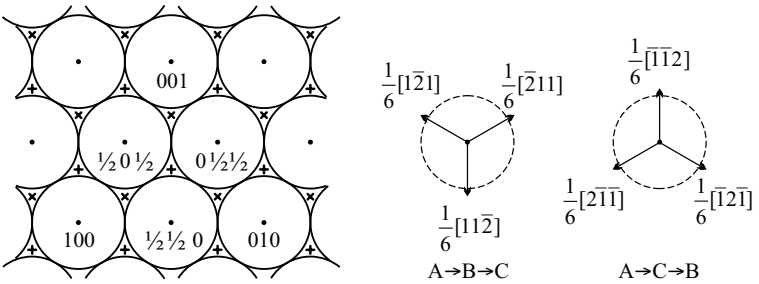


Fig. 9.16. The (111) plane through atoms at points 100, 010, 001 in a face-centered cubic lattice. Symbols × and + indicate the centers of atoms in the next (B-type) and previous (C-type) layers. The right-hand side shows the vectors that take different layers into each other

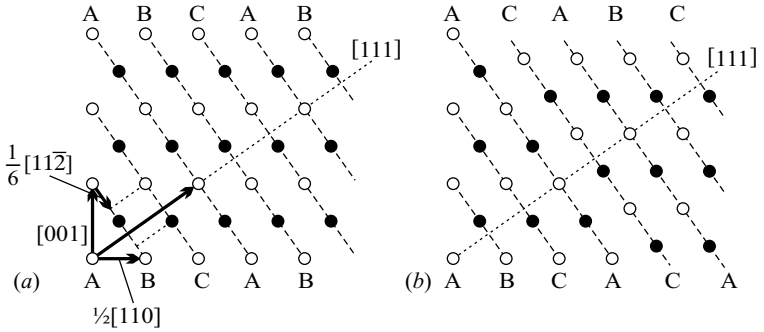


Fig. 9.17. Atoms in the $(1\bar{1}0)$ plane through the origin and in the neighboring planes in a face-centered cubic lattice: (a) perfect crystal; (b) crystal with a stacking fault

In another stacking fault a C-type layer appears between neighboring A- and B-type layers, resulting in the sequence $\ldots ABCACBCABC\ldots$. This corresponds to the translation of a part of the crystal along the (111) plane by $\frac{1}{6}[\bar{1}\bar{1}2]$, $\frac{1}{6}[2\bar{1}\bar{1}]$, or $\frac{1}{6}[\bar{1}2\bar{1}]$.

Slipping layers cannot give rise to a particular stacking fault that may easily occur during crystals growth: starting with the defect layer the stacking sequence is reversed, leading to the structure $\ldots ABCACBACBA\ldots$.

9.3.2 Partial Dislocations

The question naturally arises: What happens in a face-centered cubic lattice if the imperfect stacking of (111) planes (an A-type layer is followed by a C- rather than a B-type layer) does not extend over the entire crystal only half of it? To obtain such a defect in the Volterra construction, one has to cut the sample halfway above an A-type layer and then displace the layers above it

by $\frac{1}{6}[11\bar{2}]$, $\frac{1}{6}[\bar{2}11]$, or $\frac{1}{6}[1\bar{2}1]$. The new arrangement of the atoms in the layers below and above the cut plane is shown in Fig. 9.18.

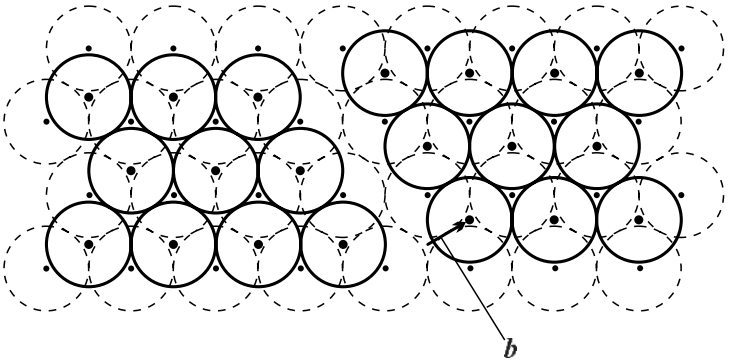


Fig. 9.18. Atomic arrangement in the (111) plane of an fcc lattice above an A-type layer, when the neighboring B-type layer is cut and the part to the right of the cutting edge $[01\bar{1}]$ is displaced perpendicularly to it, through $\boldsymbol{b} = \frac{1}{6}[\bar{2}11]$

The same atomic arrangement is shown in Fig. 9.19, projected to the plane $(01\bar{1})$ spanned by vectors $[\bar{2}11]$ and $[111]$. This view shows even more clearly that the slipped and unslipped parts of the layers above the cut plane are separated by an empty region. This can be filled by inserting two atomic layers, as shown in the figure. Because of the finite size of atomic spheres neighboring atoms are slightly displaced to make room for the inserted layers.

Since the displacement vector is not a lattice vector, the irregularity in the atomic arrangement is not limited to the vicinity of the cut edge (as for dislocations) but extends over the whole plane. Nevertheless the edge of the cut plays a privileged role here, too; in this arrangement it bears the name *Shockley partial dislocation*[9] (or just Shockley partial).

To understand the origin of the name *partial dislocation*, consider a system with an A-type layer and above it a C-type layer that is cut in the direction $[01\bar{1}]$ and then the part on the right is displaced through $\frac{1}{6}[\bar{1}2\bar{1}]$. As shown in Fig. 9.20, the right-hand side of the top layer is a B-type layer.

Now consider a system in which a part of the B-type layer is displaced in the manner shown in Figs. 9.18 and 9.19 – giving rise to a stacking fault that extends over a half plane and is bordered by a Shockley partial dislocation –, and then a portion of the obtained C-type layer is again displaced, this time by $\frac{1}{6}[\bar{1}2\bar{1}]$, leading to another B-type (half) layer. The atomic arrangement in the $(01\bar{1})$ plane is shown in Fig. 9.21; deformations are taken into account.

Atomic layers are perfectly in registry on the left and right parts of the sample; misfit is limited to the finite region between the two cut lines. As

[9] W. B. Shockley, 1948.

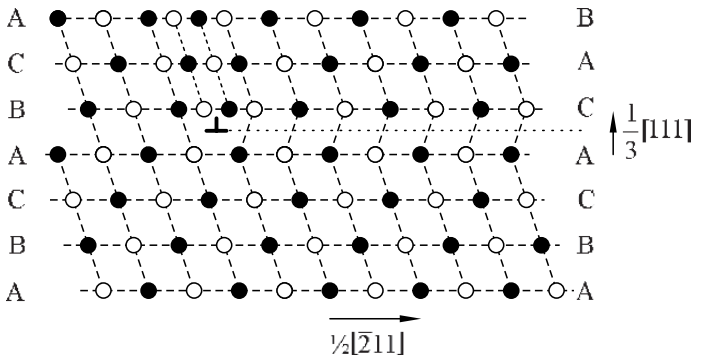


Fig. 9.19. Atomic arrangement around a Shockley partial dislocation. Atoms marked by empty circles lie in a $(01\bar{1})$ plane, while those marked by full circles lie in the parallel plane in front or behind, separated by a quarter of the face diagonal $[01\bar{1}]$

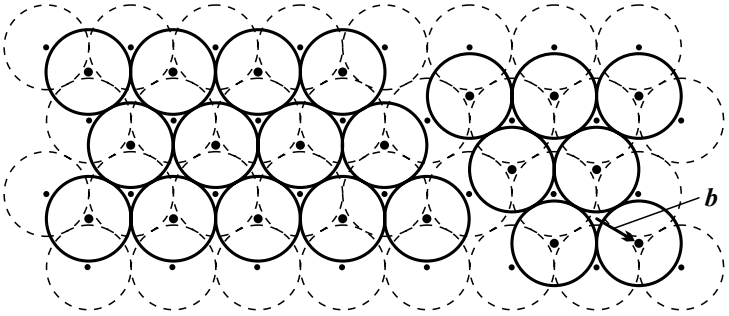


Fig. 9.20. Atomic arrangement in the (111) plane of an fcc lattice above an A-type layer, when the neighboring C-type layer is cut and the right part is displaced through $\boldsymbol{b} = \frac{1}{6}[\bar{2}11]$

$$\tfrac{1}{6}[\bar{2}11] + \tfrac{1}{6}[\bar{1}2\bar{1}] = \tfrac{1}{2}[\bar{1}10]\,, \tag{9.3.1}$$

and since $\frac{1}{2}[\bar{1}10]$ is a primitive vector in a face-centered cubic lattice, when the chosen Burgers circuit encloses the whole defect region with both cut lines, the topology is the same as if in its interior there were an ordinary dislocation characterized by the Burgers vector $\frac{1}{2}[\bar{1}10]$.

Even when such a dislocation is formed originally, it is energetically favorable to have substantial atomic displacements not only in the immediate vicinity of the dislocation line but over a finite two-dimensional region. This extended defect can also be considered as if the dislocation had split into two partial dislocations. The region with imperfect stacking is bounded by these partial dislocations.

For Shockley partials the displacement vector – the Burgers vector – (in this particular case $\frac{1}{6}[\bar{2}11]$ or $\frac{1}{6}[\bar{1}2\bar{1}]$) is parallel to the plane of the stacking

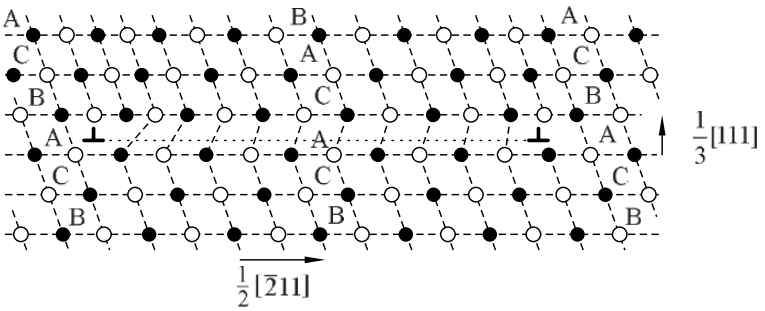


Fig. 9.21. Projection of a finite stacking fault bordered by partial dislocations in a face-centered cubic lattice on the $(01\bar{1})$ plane. Atoms marked by empty circles lie in a $(01\bar{1})$ plane, while those marked by full circles lie in the parallel plane in front or behind, separated by a quarter of the face diagonal $[01\bar{1}]$

fault. When the displacement vector is perpendicular to the plane of the stacking fault, we speak of *Frank partial dislocations* or *Frank partials*.[10] They are shown in Fig. 9.22.

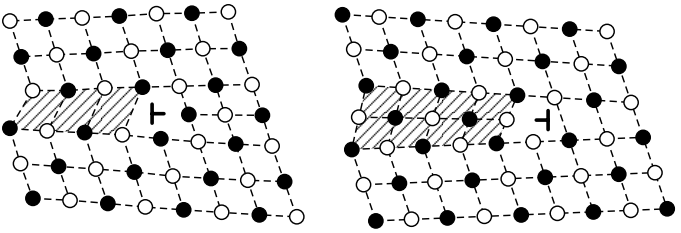


Fig. 9.22. Positive and negative Frank partial dislocations in a face-centered cubic crystal

To obtain such a defect in the Volterra construction, one has to insert or remove an atomic layer along the cut plane. Since this will necessarily disturb the matching of A, B, and C-type layers, it will give rise to an extended planar defect. As the Burgers vector of the Frank partial dislocation is $\frac{1}{3}[111]$, the relation

$$\tfrac{1}{6}[\bar{2}11] + \tfrac{1}{3}[111] = \tfrac{1}{2}[011] \tag{9.3.2}$$

implies that the complex defect of a Frank and a relatively close Schokley partial is topologically equivalent to a dislocation. Thus in face-centered cubic lattices one may observe stacking faults that are bordered by two different types of partial dislocation. The converse is also true. Rearrangement of the terms in the previous equation shows that a dislocation with a Burgers vector

[10] F. C. Frank, 1949.

$\frac{1}{2}[011]$ and a nearby Shockley partial with Burgers vector $\frac{1}{6}[2\bar{1}\bar{1}]$ are together equivalent to a Frank partial dislocation.

9.3.3 Low-Angle Grain Boundaries

In crystal-growth processes individual grains start to grow independently, and the orientation of their crystallographic axes is also uncorrelated. Due to the misfit in the atomic arrangement, a finite surface energy is associated with boundaries between grains in polycrystalline materials. It is energetically favorable to have low-angle grain boundaries with a misorientation angle of not more than a few (at most 10) degrees between the crystallographic axes of adjacent crystallites. When a crystal is made up of slightly misoriented micron-sized crystallites, we speak of a *mosaic structure*. Figure 9.23 shows two manifestations of low-angle grain boundaries.

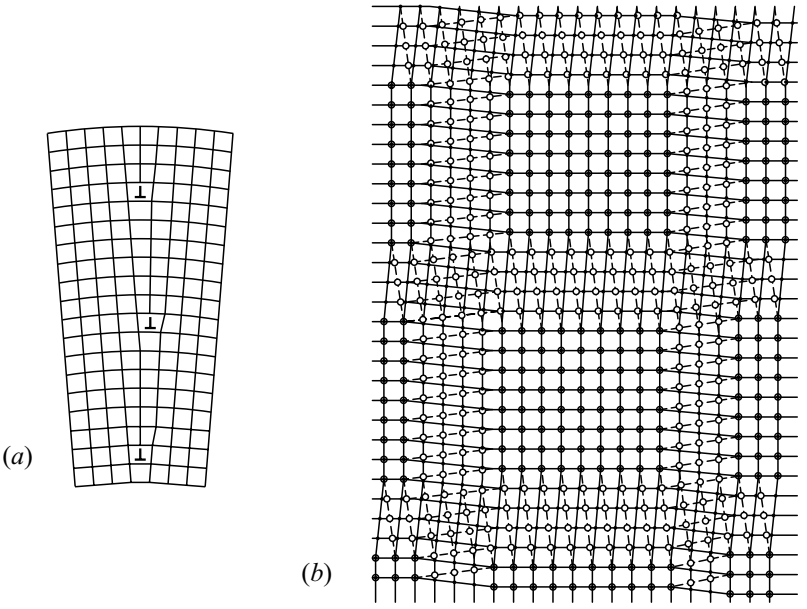


Fig. 9.23. Grain boundaries in mosaic structures: (*a*) tilt grain boundary; (*b*) twist grain boundary

In the first case the boundary is perpendicular to the plane of the figure; in the second it is parallel to it, which is why the arrangement of atoms is shown in two layers on top of each other. In either case, crystallographic axes are slightly misoriented on the two sides of the boundary. The rotation axis associated with the misorientation is in the boundary plane for *tilt grain boundaries* and perpendicular to it for *twist grain boundaries*. By taking a closer look at

the grain boundary, nets of edge and screw dislocations are observed in the two cases. In larger-angle grain boundaries the dislocation cores are so close together that one can no longer speak of individual dislocations.

9.3.4 Coincident-Site-Lattice and Twin Boundaries

When the crystallographic orientation of two neighboring grains differs appreciably, the energy of the grain boundary is higher. It may nevertheless happen, especially during recrystallization, that adjacent crystallites become somewhat reoriented, and give rise to relatively low-energy boundaries, even though the angle is not small. This requires some special relationship between the orientations. One type of high-angle grain boundary corresponds to the situation in which, despite their different orientations, the two grains can be considered in a loose sense as each other's continuation. As an illustration consider the case shown in Fig. 9.24: a special high-angle tilt and twist grain boundary between two grains of a cubic (square) crystal. It is clearly seen that when the tilt or twist angle is 36.9°, every fifth atom would coincide if the two grains were continued across the boundary. These lattice points constitute a *coincident-site lattice* (CSL). Since the density of coincident sites is one-fifth, it is called a $\Sigma 5$ CSL.

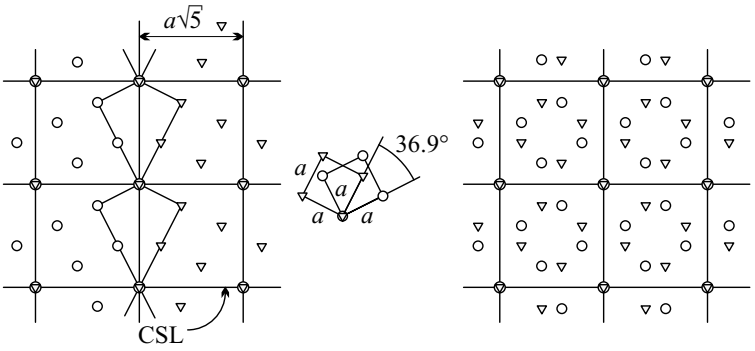


Fig. 9.24. The $\Sigma 5$ coincident-site lattice of two grains in a cubic crystal tilted/twisted by 36.9°. The left side of the figure shows a tilt boundary, and the right one a twist boundary

Of special interest is the case where the close-packed planes of two grains of a face-centered cubic crystal are tilted by 38.2° around the [111] direction. Figure 9.25 shows the atomic arrangement in the vicinity of such a grain boundary. It is clearly seen that in spite of the grain boundary, a regular lattice is formed by every seventh atom in the close-packed plane. Such a grain boundary is called a *coincident-site-lattice boundary* (CSL boundary) – in this particular case a $\Sigma 7$ boundary. In this geometry there is hardly any

empty space between the two grains, and when atoms in the grain boundary are slightly displaced each atom may have sixfold coordination in the plane, thus the energy increment due to the presence of a grain boundary is reduced.

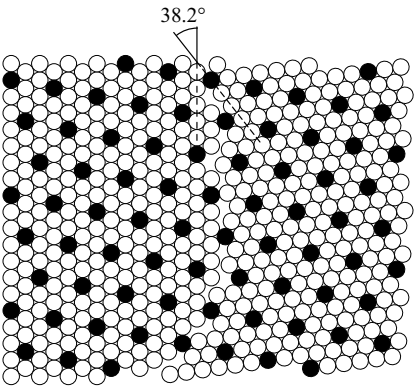


Fig. 9.25. Coincident-site lattice in the close-packed plane of a face-centered cubic crystal formed between two grains with a relative orientation of 38.2°. The atoms marked by full circles exhibit regular crystalline order

Another type of low-energy grain boundary may develop during the crystal-growth process. It corresponds to the particular situation when crystallographic axes in the two grains are taken into each other by a geometrical transformation (reflection in a plane, rotation through 180°, or inversion). In such cases we speak of a *twin crystal* and a *twin boundary*. A simple possibility was mentioned in the discussion of stacking faults in face-centered cubic lattices: starting with a (111) plane, the stacking sequence of close-packed hexagonal planes may be reversed, giving rise to an $\ldots ABC\,ABC\,ACBACBA\ldots$-type structure. Crystallographic axes on the two sides of the plane are each other's mirror images. Another situation is illustrated in Fig. 9.26, where the atomic arrangement is projected on the plane with Miller indices $(1\bar{1}0)$ in a crystal with a twin boundary at a $(11\bar{2})$ plane. The crystallographic axes of the two grains are each other's mirror images in the plane.

9.3.5 Antiphase Boundaries

Finally we mention another possible atomic arrangement for grains grown in different parts of the sample that gives rise to a special boundary. Consider a two-component ordered alloy of composition AB, e.g. CuAu with an $L1_0$ structure. Atoms of either component occupy the vertices of the tetragonal primitive cells while atoms of the other component are at the centers of the cells; obviously either sublattice can be assigned to either component. Figure 9.27 shows in a double volume unit cell that, corresponding to the two

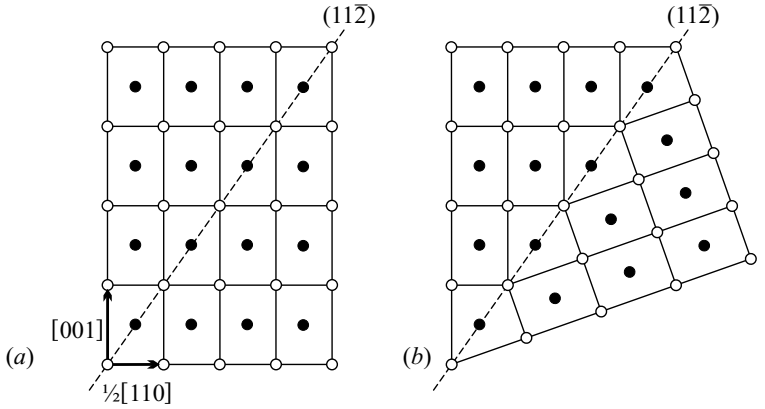


Fig. 9.26. (*a*) Atomic arrangement in a face-centered cubic crystal in the plane perpendicular to the direction [1$\bar{1}$0]. (*b*) Atomic arrangement in a twin crystal with a twin boundary at a (11$\bar{2}$) plane. Atoms in the (1$\bar{1}$0) plane through the origin are marked by empty circles; atoms in front of and behind this plane are marked by full circles

possibilities, atoms of the first component are at the vertices and base centers, while those of the second component are at the side centers – or vice versa.

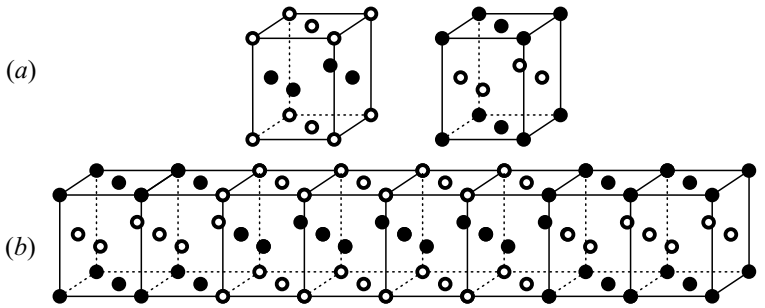


Fig. 9.27. (*a*) Two possible atomic arrangements in an ordered CuAu alloy. (*b*) Antiphase domain

When an alloy is cooled from its high-temperature disordered state, it may occur during the disorder–order transition that the same sublattice is occupied by the first component in one part of the sample and the second component in another. In such cases an *antiphase domain* is said to have been formed. The interface between the two regions is called an *antiphase boundary*.

9.4 Volume Defects

If the composition – and therefore the atomic arrangement – inside a small nevertheless macroscopic region of a polycrystalline material is different from the rest of the sample, we speak of a *volume* or *bulk defect*. When the concentration of an impurity exceeds the maximum solubility in the crystal, three-dimensional regions can be formed where the excess impurities cluster together. The chemical composition of these three-dimensional defects, called *precipitates*, is therefore different from the host lattice. Similar defects may appear in nonstoichiometric, multicomponent materials as well.

During crystal growth or fabrication processes it may also happen that small but macroscopic sized empty regions, *voids* remain in the sample. These can be considered as if a macroscopic number of vacancies were present in a block.

Further Reading

1. H. G. van Bueren, *Imperfections in Crystals*, North-Holland Publishing Co., Amsterdam (1960).
2. J. Friedel, *Dislocations*, Pergamon Press, Oxford (1964).
3. J. P. Hirth and J. Lothe, *Theory of Dislocations*, McGraw-Hill Book Company, New York (1968).
4. D. Hull and D. J. Bacon, *Introduction to Dislocations*, Pergamon Press, Oxford (1984).
5. M. Kléman, *Points, Lines and Walls: In Liquid Crystals, Magnetic Systems and Various Ordered Media*, John Wiley & Sons, Chichester (1983).
6. F. R. N. Nabarro, *Theory of Crystal Dislocations*, Oxford University Press, Oxford (1967).
7. B. K. Vainshtein, V. M. Fridkin, and V. L. Indenbom, *Modern Crystallography*, Volume 2; *Structure of Crystals*, Third edition, Springer-Verlag, Berlin (2000).

10

The Structure of Noncrystalline Solids

We started our investigation into solid-state physics with the statement that in thermal equilibrium at low temperatures atoms in a solid are expected to be arranged in a regular periodic array, and so solids are characterized by a long-range positional order. However, it has been known for a long time that there exist certain materials that can be considered as solids from a mechanical point of view although the atomic arrangement in them is disordered as in liquids. They only exhibit short-range order, if any. It has also been shown recently that certain solids may have a state with some kind of long-range order even though atoms are not arranged periodically. In the present chapter we shall discuss the structural characteristics of such materials.

10.1 The Structure of Amorphous Materials

In solids prepared from melts by quenching, cooling may be so rapid that it leaves no time for nucleation and crystal growth, and so the equilibrium crystalline phase is not reached. Atoms freeze into a disordered, thermodynamically metastable state that is similar to the liquid phase or to a glassy (vitreous) state.

Similar structures are formed when atoms are evaporated on a surface, and randomly arriving atoms are stacked on top of each other in a disordered way. Solid, mechanically rigid materials in which atoms are arranged randomly are called amorphous.[1] No long-range positional order exists inside them but they may exhibit short-range order.

10.1.1 Models of Topological Disorder

Two simple models for the structure of disordered materials are worth mentioning. In the first one atoms considered as rigid spheres are packed randomly

[1] Amorphous metallic materials produced by quenching alloys that are close to their eutectic points are also called *metallic glasses*.

but as tightly as possible. This *dense random packing model* is in fact the extension of the Bernal model of liquids to the amorphous, glassy state. This is illustrated in two dimensions in Fig. 2.2, where the plane is covered with randomly arranged but touching circles.

This can be used to model the structure of amorphous materials in which atoms have a repulsive core but otherwise tend to approach each other as much as possible. This is usually the case for metals – which explains why hexagonal close-packed and face-centered cubic structures occur so often in metals. In amorphous metals short-range order is due to this tendency for close packing. As we have seen, in close-packed fcc and hcp crystals each atomic sphere touches 12 neighbors, giving a packing fraction of almost 75%. In randomly packed structures – even when spheres touch as many others as possible – space filling is less efficient than in regular lattices. Empirical evidence shows that the packing fraction is at most 64%; on the average each sphere touches 8 to 9 others. A somewhat better space filling is obtained in numerical simulations when the randomly placed but touching atomic spheres are allowed small displacements toward local energy minima in their respective potential fields due to the other atoms (i.e., along the direction of the gradient of the full potential felt by the atom).

The other model, which leads to much less efficient space filling, is applicable to covalently bonded systems. In Chapters 4 and 7 we saw that covalent bonds are highly directional, and fix the relative orientation of nearest-neighbor atoms. A prime example was the diamond structure in which each atom is surrounded by four neighbors in a unique tetrahedral arrangement. The crystal structure can be considered as the network of such bonds. Relative orientation of the tetrahedra plays a key role in forming the structure. The wurtzite and sphalerite structures of zinc sulfide (ZnS) differ precisely in the relative orientation of the tetrahedra. Similar tetrahedral geometry is observed in quartz (SiO_2). Around each silicon atom oxygen atoms are placed in the directions of the four vertices of a tetrahedron. Joined at their vertices, these tetrahedra make up a three-dimensional network. Since neighboring tetrahedra may easily rotate with respect to one another, quartz often appears in an amorphous state.

In Fig. 7.16 showing diamond and sphalerite structures tetrahedral bonds are seen to form hexagonal rings. When the tetrahedra are rotated, there is a finite probability for a ring to contain seven or five rather than six atoms. A two-dimensional illustration is shown in Fig. 10.1; here each atom is bonded to three others in the plane. The regular array would be a honeycomb, however rings of five and seven also appear in the disordered structure.

As for covalent bonds showing local tetrahedral symmetry in three-dimensional space, even when the local structure is maintained as much as possible, the appearance of disorder necessarily changes the topology: firstly, pentagonal and heptagonal rings appear, and secondly, some atoms form bonds with three neighbors only. Electron states that do not participate in

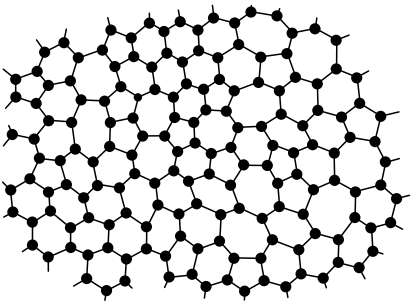


Fig. 10.1. Amorphous structures with a topological disorder may contain pentagonal and heptagonal rings as well

bonding can be considered to form *dangling bonds*. They strongly affect the properties of covalently bonded amorphous materials.

10.1.2 Analysis of the Short-Range Order

Whichever previous model of disordered systems is adopted, the relative position of nearest neighbors around each atom is similar to that observed in crystalline structures. However, on larger scales the correlation between atomic positions is completely washed out. The presence of a short-range order is indicated by the peaks of the radial distribution function (2.1.13) at first- and second-neighbor distances. Below we shall demonstrate that the radial distribution function lends itself to simple experimental determination in terms of the angular distribution of the intensity of elastically scattered X-ray or neutron beams.

In Section 8.1 on the theory of diffraction we saw that if the scattering amplitude of an atom at position $\boldsymbol{R}_m$ is f_m then the amplitude of the beam scattered from incident direction $\boldsymbol{k}$ into the direction of $\boldsymbol{k}'$ is

$$A_{\boldsymbol{K}} = \sum_m f_m(\boldsymbol{K}) \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{R}_m} \,, \tag{10.1.1}$$

where $\boldsymbol{K} = \boldsymbol{k} - \boldsymbol{k}'$. The intensity of the scattered beam is proportional to

$$|A_{\boldsymbol{K}}|^2 = \sum_{m,n} f_m(\boldsymbol{K}) f_n^*(\boldsymbol{K}) \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot(\boldsymbol{R}_m - \boldsymbol{R}_n)}. \tag{10.1.2}$$

When the incident beam is scattered by a system composed of identical atoms, the expression for intensity contains the Fourier transform

$$\Gamma(\boldsymbol{K}) = \frac{1}{N} \sum_{m,n} \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot(\boldsymbol{R}_m - \boldsymbol{R}_n)} \tag{10.1.3}$$

of the correlation function defined in (2.1.16). Using (2.1.19) this can be related to the structure factor $S(\boldsymbol{K})$, and then, making use of (2.1.22), the

pair-correlation function $c(\boldsymbol{r})$ and the distribution function $g(\boldsymbol{r})$ can be derived. It should be emphasized once again that these relations connecting the elastic scattering cross section to the structure factor and thence to the distribution function are not limited to crystalline materials but are valid without restriction.

In the previous chapter we saw that in crystalline materials sharp diffraction peaks appear at vectors $\boldsymbol{K} = \boldsymbol{G}$ that satisfy the Bragg condition. This is not the case in noncrystalline solids. When the atomic positions are randomly distributed, rays scattered by different atoms interfere destructively because of the random phase difference, so $S(\boldsymbol{K}) = 1$ for every $\boldsymbol{K}$ except $\boldsymbol{K} = 0$. Consequently $g(\boldsymbol{r}) = 1$, that is, the pair-correlation function is identically zero. In amorphous materials, where short-range order is present, $S(\boldsymbol{K})$ is not constant. Since amorphous samples are isotropic, spatial variations of the pair distribution function – i.e., the radial distribution function – can be determined from the variations of the elastically scattered beam using (2.1.24). Figure 10.2 shows the structure factor measured by neutron diffraction and the derived radial distribution function for amorphous silicon.

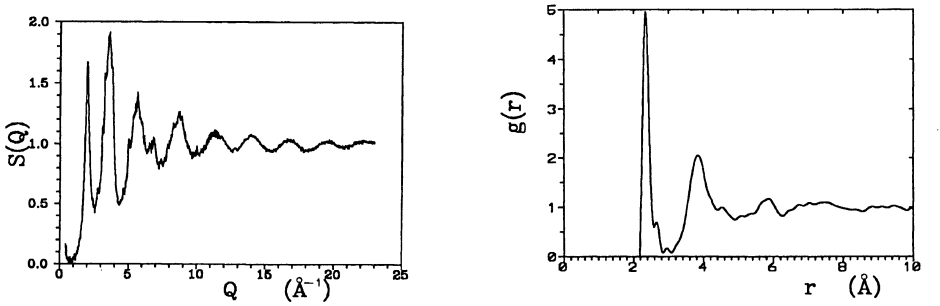


Fig. 10.2. Structure factor of amorphous silicon measured by neutron diffraction, and the derived radial distribution function. [S. Kugler et al., *Phys. Rev. B* **48**, 7685 (1993)]

Smeared-out peaks in the structure factor may be interpreted as the broadened counterparts of the sharp Bragg peaks observed in crystals. This occurs because only short-ranged order is present and correlations exist only among first, second, and perhaps third neighbors. This closely parallels the situation in liquids, as discussed in Chapter 2.

For multicomponent amorphous materials diffraction measurements reveal not only the total distribution function but also the so-called partial distribution functions. Consider, for example, a two-component system made up of N_A atoms of type A and N_B atoms of type B, both distributed randomly. The concentrations are

$$c_i = \frac{N_i}{N_\mathrm{A} + N_\mathrm{B}} \qquad i = \mathrm{A}, \mathrm{B}\,. \tag{10.1.4}$$

If scattering amplitudes are denoted by f_{A} and f_{B} and atomic positions by $\boldsymbol{R}_m^{\mathrm{A}}$ and $\boldsymbol{R}_n^{\mathrm{B}}$ then the amplitude of the scattered beam in the direction specified by the scattering vector $\boldsymbol{K} = \boldsymbol{k} - \boldsymbol{k}'$ is the weighted sum of the amplitudes of the rays scattered by individual atoms. The intensity is therefore

$$I(\boldsymbol{K}) = I_0 \left| \sum_m f_{\mathrm{A}} \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{R}_m^{\mathrm{A}}} + \sum_n f_{\mathrm{B}} \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{R}_n^{\mathrm{B}}} \right|^2 . \tag{10.1.5}$$

Now we introduce the quantities

$$\Gamma_{ij}(\boldsymbol{r}) = \frac{1}{N} \left\langle \sum_{m=1}^{N_i} \sum_{n=1}^{N_j} \delta(\boldsymbol{r} - \boldsymbol{R}_m^i + \boldsymbol{R}_n^j) \right\rangle , \tag{10.1.6}$$

where N is the total number of atoms. Their Fourier transforms appear in the intensity

$$I(\boldsymbol{K}) = I_0 N \Gamma(\boldsymbol{K}) , \tag{10.1.7}$$

where, if the scattering amplitudes are chosen real,

$$\Gamma(\boldsymbol{K}) = f_{\mathrm{A}}^2 \Gamma_{\mathrm{AA}}(\boldsymbol{K}) + f_{\mathrm{B}}^2 \Gamma_{\mathrm{BB}}(\boldsymbol{K}) + 2 f_{\mathrm{A}} f_{\mathrm{B}} \Gamma_{\mathrm{AB}}(\boldsymbol{K}) . \tag{10.1.8}$$

The contribution $\boldsymbol{K} = 0$ of forward scattering is customarily separated off from $\Gamma_{ij}(\boldsymbol{K})$; the remaining part defines the partial structure factor $S_{ij}(\boldsymbol{K})$

$$\Gamma_{ij}(\boldsymbol{K}) = c_i c_j N \delta_{\boldsymbol{K},0} + c_i c_j S_{ij}(\boldsymbol{K}) . \tag{10.1.9}$$

Substituting this into (10.1.8),

$$\begin{aligned} \Gamma(\boldsymbol{K}) = {} & N \delta_{\boldsymbol{K},0} \left[f_{\mathrm{A}}^2 c_{\mathrm{A}}^2 + f_{\mathrm{B}}^2 c_{\mathrm{B}}^2 + 2 f_{\mathrm{A}} f_{\mathrm{B}} c_{\mathrm{A}} c_{\mathrm{B}} \right] \\ & + f_{\mathrm{A}}^2 c_{\mathrm{A}}^2 S_{\mathrm{AA}}(\boldsymbol{K}) + f_{\mathrm{B}}^2 c_{\mathrm{B}}^2 S_{\mathrm{BB}}(\boldsymbol{K}) + 2 f_{\mathrm{A}} f_{\mathrm{B}} c_{\mathrm{A}} c_{\mathrm{B}} S_{\mathrm{AB}}(\boldsymbol{K}) . \end{aligned} \tag{10.1.10}$$

Direct scattering is seen to be proportional to $(f_{\mathrm{A}} c_{\mathrm{A}} + f_{\mathrm{B}} c_{\mathrm{B}})^2$ – as if the medium were monatomic, with an averaged scattering amplitude.

Partial structure factors are related to partial pair-correlation functions. Similarly to (2.1.1), up to a normalization factor $g_{ij}(\boldsymbol{r})$ is defined as the probability for finding an atom of the jth type in an elementary volume $\mathrm{d}\boldsymbol{r}$ around the point $\boldsymbol{r}$ when an atom of the ith type is located at the origin:

$$g_{ij}(\boldsymbol{r}) = \frac{V}{N_i N_j} \left\langle \sum_{m=1}^{N_i} \sum_{n=1}^{N_j} \delta(\boldsymbol{r} - \boldsymbol{R}_m^i + \boldsymbol{R}_n^j) \right\rangle . \tag{10.1.11}$$

In the expressions for g_{AA} and g_{BB} the terms $m = n$ are excluded.

Along the same lines as in Chapter 2 it can be shown that

$$\Gamma_{ij}(\boldsymbol{r}) = \delta_{ij}\, c_i\, \delta(\boldsymbol{r}) + n c_i c_j g_{ij}(\boldsymbol{r}) , \tag{10.1.12}$$

where $n = N/V$; in terms of the Fourier transforms

$$\Gamma_{ij}(\boldsymbol{K}) = \delta_{ij}\, c_i + n c_i c_j g_{ij}(\boldsymbol{K})\,. \tag{10.1.13}$$

Comparing this form with (10.1.9), making use of (2.1.20), and taking into account that amorphous materials are isotropic, one obtains

$$S_{ij}(K) = \delta_{ij}\frac{1}{c_i} + \frac{N}{V}\int\limits_0^\infty 4\pi r^2\,[g_{ij}(r) - 1]\,\frac{\sin Kr}{Kr}\,\mathrm{d}r\,, \tag{10.1.14}$$

and

$$g_{ij}(r) = 1 + \frac{V}{8\pi^3 N}\int\limits_0^\infty 4\pi K^2\left[S_{ij}(K) - \delta_{ij}\frac{1}{c_i}\right]\frac{\sin Kr}{Kr}\,\mathrm{d}K\,, \tag{10.1.15}$$

which are the generalizations of (2.1.23) and (2.1.24).

In neutron scattering the scattering amplitudes depend strongly on which particular isotope is used. Performing diffraction measurements for various isotopic compositions, the partial structure factors, and from them the partial radial distribution functions may be determined. Such a set of experimental results is shown in Fig. 10.3.

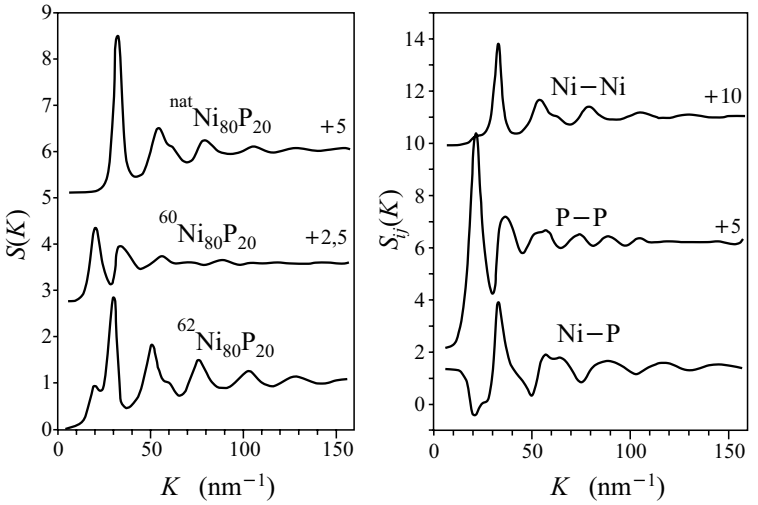


Fig. 10.3. Structure factors measured for various isotopic compositions of a two-component amorphous material, $Ni_{80}P_{20}$, and the derived partial structure factors. For better visibility, curves are shifted vertically. [P. Lamparter and S. Steeb, *Proc. of the Fifth Int. Conf. on Rapidly Quenched Metals*, p. 459 (1984)]

Besides the structure factors measured for three different isotopic compositions, the figure also shows the partial structure factors derived from

them. Their Fourier transform yields the pair-correlation functions shown in Fig. 10.4, more precisely the quantities

$$G_{ij}(r) = 4\pi r[g_{ij}(r) - 1]\,. \tag{10.1.16}$$

As the figure shows, nearest neighbor Ni–Ni and Ni–P distances are practically equal, while nearest-neighbor P–P distances are larger than these. The arrangement of the two kinds of atom is therefore not perfectly random. Short-range order appears because phosphorus atoms are preferentially surrounded by nickel atoms.

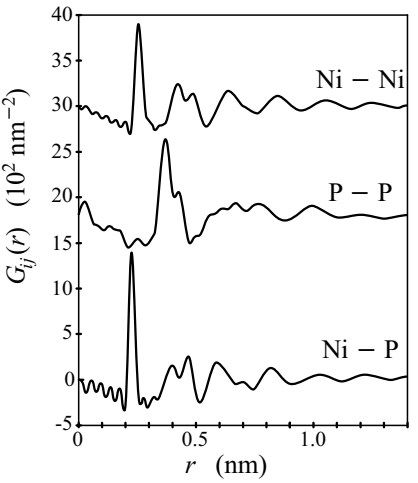


Fig. 10.4. Partial distribution functions in amorphous $Ni_{80}P_{20}$. [P. Lamparter and S. Steeb, *Proc. of the Fifth Int. Conf. on Rapidly Quenched Metals*, p. 459 (1984)]

An efficient method for determining the local environment in disordered systems is EXAFS spectroscopy as the fine structure of the absorption edge is due to the interference of rays scattered by nearest neighbors. Figure 10.5 shows that the EXAFS spectrum of amorphous GeO_2 has the same fine structure as GeO_2 in its hexagonal crystalline form, and both are strikingly different from the spectrum observed in samples of tetragonal symmetry. This leads to the conclusion that locally, on atomic scales the arrangement of atoms in the amorphous state is very similar to that in the hexagonal crystalline phase.

10.2 Quasiperiodic Structures

In contrast to perfectly periodic order in crystals, amorphous materials may exhibit short-range order at best. In the past decades it has been revealed that nature offers a wider variety of solid structures than that presented in the

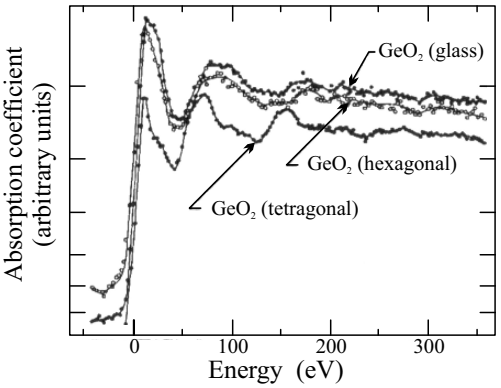


Fig. 10.5. Fine structure of the absorption edge in amorphous and crystalline GeO_2. [W. F. Nelson et al., *Phys. Rev.* **127**, 2025 (1962)]

foregoing. There exist materials that do not possess discrete translation symmetries but nonetheless exhibit long-range order of some kind. Before turning to the investigation of such physical systems, some mathematical remarks are due.

10.2.1 Periodic and Quasiperiodic Functions

In Section 8.1 on the theory of diffraction we saw that in systems composed of identical atoms the amplitude of the scattered beam is proportional to the structure factor, which is related to the Fourier transform of the spatial distribution of scatterers. Below we shall examine the differences from this viewpoint between periodic, aperiodic (nonperiodic), almost periodic, and quasiperiodic functions. We shall use single-variable functions, however the obtained results are straightforward to generalize to the case of three spatial dimensions.

If a function $f(x)$ is periodic with a period L – that is, $f(x+L) = f(x)$ for all values of x – then it can be expanded in a Fourier series as

$$f(x) = \sum_{n=-\infty}^{\infty} f_n \mathrm{e}^{2\pi \mathrm{i} n x/L}, \tag{10.2.1}$$

where n is an integer and

$$f_n = \frac{1}{L} \int_{-L/2}^{L/2} \mathrm{e}^{-2\pi \mathrm{i} n x/L} f(x)\, \mathrm{d}x\,. \tag{10.2.2}$$

The Fourier spectrum of a periodic function is therefore discrete. This corresponds to a sequence of equidistant Bragg peaks in the diffraction pattern of a one-dimensional chain.

Changing to the variable $k = 2\pi n/L$ and taking the limit $L \to \infty$ opens the way to representing any nonperiodic function by a Fourier integral:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(k) \mathrm{e}^{\mathrm{i}kx} \, \mathrm{d}k \,, \tag{10.2.3}$$

where

$$\hat{f}(k) = \int_{-\infty}^{\infty} f(x) \mathrm{e}^{-\mathrm{i}kx} \, \mathrm{d}x \,. \tag{10.2.4}$$

In this case the Fourier spectrum is continuous. This is typical of disordered systems.

There are, however, nonperiodic functions whose Fourier spectrum is discrete. Suppose that the Fourier transform of $f(x)$ takes finite values on an infinite sequence λ_n of real numbers, that is its Fourier representation may be written as

$$f(x) = \sum_n \hat{f}(\lambda_n) \exp(2\pi \mathrm{i} \lambda_n x) \,. \tag{10.2.5}$$

Such a function is called *almost periodic*: although it is not perfectly periodic in space, for every ε there exists a finite distance t such that

$$|f(x+t) - f(x)| \leq \varepsilon \tag{10.2.6}$$

for all values of x.

Quasiperiodic functions constitute a special class of almost periodic functions. A function is called quasiperiodic if the infinitely many λ_n of its spectrum can be written as integer linear combinations of a finite number of irrational numbers α_l $(l = 1, 2, \ldots, N)$:

$$\lambda_n = \sum_{l=1}^{N} n_l \alpha_l \,, \tag{10.2.7}$$

and the coefficients n_l can take any integer values. In what follows, the label n will continue to refer to the set of coefficients $n_1, n_2, \ldots, n_N$. Using this notation, the Fourier representation reads

$$f(x) = \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} \cdots \sum_{n_N=-\infty}^{\infty} f_n \exp\left[2\pi \mathrm{i} \Big(\sum_{l=1}^{N} \alpha_l n_l \Big) x\right] . \tag{10.2.8}$$

Next we shall show that this function can be obtained from a function in N variables that is perfectly periodic in N-dimensional space when the values $x_1, x_2, \ldots, x_N$ lie along a line. Consider the periodic function

$$F(x_1, x_2, \ldots, x_N) = \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} \cdots \sum_{n_N=-\infty}^{\infty} f_n \exp\left[2\pi \mathrm{i} \Big(\sum_{l=1}^{N} n_l x_l \Big)\right] \tag{10.2.9}$$

defined in N-space. The quasiperiodic function $f(x)$ can be obtained by restricting the previous expression to the line

$$x_1 = \alpha_1 x, \quad x_2 = \alpha_2 x, \ldots, \quad x_N = \alpha_N x\,, \tag{10.2.10}$$

that is,

$$f(x) = F(\alpha_1 x, \alpha_2 x, \ldots, \alpha_N x)\,. \tag{10.2.11}$$

Generalization of the above formulation to functions defined in three-dimensional space is straightforward. It has already been mentioned that if a function $f(\boldsymbol{r})$ is periodic in the lattice generated by the primitive translation vectors $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$ then its Fourier series contains only components associated with vectors

$$\boldsymbol{k} = \boldsymbol{G} = \sum_{i=1}^{3} h_i \boldsymbol{b}_i \tag{10.2.12}$$

of the reciprocal lattice, where the $\boldsymbol{b}_i$ are the primitive vectors of the reciprocal lattice and the h_i are integers. A function defined in three-dimensional space is called quasiperiodic if the vectors $\boldsymbol{k}$ in its Fourier expansion can be expressed as the linear combinations of more than three reciprocal-space vectors $\boldsymbol{q}_i$ with integer coefficients h_i,

$$\boldsymbol{k} = \sum_{i=1}^{N} h_i \boldsymbol{q}_i\,, \tag{10.2.13}$$

while the vectors $\boldsymbol{q}_i$ are irrational linear combinations of the three primitive vectors $\boldsymbol{b}_i$. Then the allowed $\boldsymbol{k}$ vectors fill the reciprocal space densely. Nevertheless, as we shall see, the amplitude is very small for most of them, and the diffraction pattern is dominated by a few discrete peaks.

10.2.2 Incommensurate Structures

Up to this point we have always spoken of periodic or nonperiodic arrangements of atoms. However, in solids the system of electrons may exhibit a periodicity different from that of the atoms, with the new wavelength determined by the properties of the electron system. We shall see examples for this in Chapters 14 and 33. It may happen that the wavelength ratio of atomic and electronic periodicities is not a rational number; in such cases the system is said to have an *incommensurate structure*. Strictly speaking periodicity is lost as the structure never repeats itself precisely because of the incommensurability of the two periods, nonetheless there exists a long-range order as it is possible to predict atomic and electronic densities precisely for any point in space.

To determine what diffraction pattern would be obtained from a material with such a structure we consider, for simplicity, a chain of atoms with a lattice constant a and suppose that the electron system shows static density variations of wavelength λ that are described by a periodic function $g(x)$:

$$g(x + p\lambda) = g(x) \tag{10.2.14}$$

for all integers p. Because of the interaction between delocalized electrons and localized atoms the latter are expected to be displaced from their equilibrium positions, and the displacement to depend on the local electron density. Assuming a linear relationship, the new equilibrium position for the nth atom is

$$na + ug(na)\,, \tag{10.2.15}$$

where the proportionality factor u that gives the modulation amplitude is determined by the interactions between electrons and atoms.

The density of atoms is given by

$$\rho(x) = \sum_{n=-\infty}^{\infty} \delta[x - na - ug(na)]. \tag{10.2.16}$$

When a beam is scattered by the system, the intensity of the diffracted beam is determined by the absolute square of the structure amplitude

$$\begin{aligned} F_K = \int_{-\infty}^{\infty} \rho(x) \mathrm{e}^{-\mathrm{i}Kx}\,\mathrm{d}x &= \sum_{n=-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}K[na+ug(na)]} \\ &= \sum_{n=-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}Kna} \mathrm{e}^{-\mathrm{i}Kug(na)} \end{aligned} \tag{10.2.17}$$

defined in analogy to (8.1.27). If $g(x)$ is a periodic function with period λ then so is $\exp[-\mathrm{i}Kug(x)]$, therefore it can be expanded into a Fourier series as

$$\mathrm{e}^{-\mathrm{i}Kug(x)} = \sum_{h=-\infty}^{\infty} c_h(Ku) \mathrm{e}^{2\pi \mathrm{i}hx/\lambda}\,, \tag{10.2.18}$$

hence

$$F_K = \sum_{n=-\infty}^{\infty} \sum_{h=-\infty}^{\infty} c_h(Ku) \mathrm{e}^{-\mathrm{i}na(K-2\pi h/\lambda)}\,. \tag{10.2.19}$$

Summation over n can be simplified exploiting (C.1.46) – that is, making use of the property that the sum over n is nonzero only if $a(K - 2\pi h/\lambda)$ is an integral multiple of 2π:

$$F_K = \sum_{h=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} c_h(Ku) \delta[k - a(K/2\pi - h/\lambda)]\,. \tag{10.2.20}$$

It is immediately seen that the structure factor is finite only for those values of K that can be written as

$$K = \frac{2\pi}{a} k + \frac{2\pi}{\lambda} h\,. \tag{10.2.21}$$

Since h and k can take arbitrary integer values, besides the Bragg peaks associated with the lattice of periodicity a and the Bragg peaks due to the density variations of the electron system of periodicity λ, further peaks appear at all possible linear combinations of the two sets of reciprocal-lattice vectors. When $\lambda \gg a$, these *satellite peaks* are close to the Bragg peaks of the original lattice. Although the examined system is one-dimensional, diffraction peaks are specified by two integers. Owing to the incommensurability of the two wavelengths the allowed Ks are dense all along the line. Nevertheless the diffraction pattern is a set of relatively well separated sharp peaks (sharp dots on a film) as amplitude is usually large only for reflections with a low index. This is well illustrated by the choice $g(x) = \sin x$. Using (C.1.50) it can be shown that the expansion coefficients appearing in (10.2.20) are the Bessel functions, which are known to decrease fairly rapidly with increasing orders.

Similar conclusions apply to the case when atomic positions in a three-dimensional lattice are subject to modulation with an incommensurate periodicity. Let the position of the jth atom in the nth primitive cell be denoted by

$$\boldsymbol{r}(n,j) = \boldsymbol{R}_n + \boldsymbol{r}_j + \boldsymbol{u}_j \sin[\boldsymbol{q}\cdot(\boldsymbol{R}_n + \boldsymbol{r}_j) + \varphi_j]\,. \tag{10.2.22}$$

If the scattering amplitude of the jth atom of the primitive cell is f_j, the structure amplitude is

$$\begin{aligned} F_{\boldsymbol{K}} &= \sum_{n,j} f_j \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}(n,j)} \\ &= \sum_{n,j} f_j \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot(\boldsymbol{R}_n+\boldsymbol{r}_j)} \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{u}_j \sin[\boldsymbol{q}\cdot(\boldsymbol{R}_n+\boldsymbol{r}_j)+\varphi_j]}\,. \end{aligned} \tag{10.2.23}$$

The second exponential term can again be expanded into a series of Bessel functions,

$$F_{\boldsymbol{K}} = \sum_{n,j} \sum_{m=-\infty}^{\infty} (-1)^m f_j \mathrm{e}^{-\mathrm{i}(\boldsymbol{K}-m\boldsymbol{q})\cdot(\boldsymbol{R}_n+\boldsymbol{r}_j)} \mathrm{e}^{\mathrm{i}m\varphi_j} J_m(\boldsymbol{K}\cdot\boldsymbol{u}_j)\,. \tag{10.2.24}$$

Summation over the vectors $\boldsymbol{R}_n$ labeling the primitive cells yields finite contributions only when $\boldsymbol{K} - m\boldsymbol{q}$ is a vector of the reciprocal lattice – that is, diffraction peaks are found in directions for which

$$\boldsymbol{K} = h\boldsymbol{b}_1 + k\boldsymbol{b}_2 + l\boldsymbol{b}_3 + m\boldsymbol{q}\,. \tag{10.2.25}$$

Diffraction peaks are now specified by four integers.

10.2.3 Experimental Observation of Quasicrystals

We saw in Chapter 6 on the symmetries of crystalline structures that translational symmetry allows only two-, three-, four-, and sixfold rotation axes. It was also asserted that fivefold symmetry is ruled out because the plane cannot

be tiled perfectly with regular pentagons. When discussing the Laue method of structural analysis in Section 8.2.3 it was mentioned that the Laue pattern had to possesses the same rotational symmetries as the crystal itself around the direction of the incident beam.

In the light of these not without reason did the discovery of D. Shechtman, I. Blech, D. Gratias, and J. W. Cahn made in 1984 cause great sensation. When studying a quenched sample of an aluminum-manganese alloy ($Al_{86}Mn_{14}$) using electron diffraction methods it was noticed that, depending on the direction of observation, the Laue patterns exhibited symmetries characteristic of two-, three- and fivefold rotational axes, as if the crystal possessed icosahedral symmetry. Such diffraction patterns are shown in Fig. 10.6.

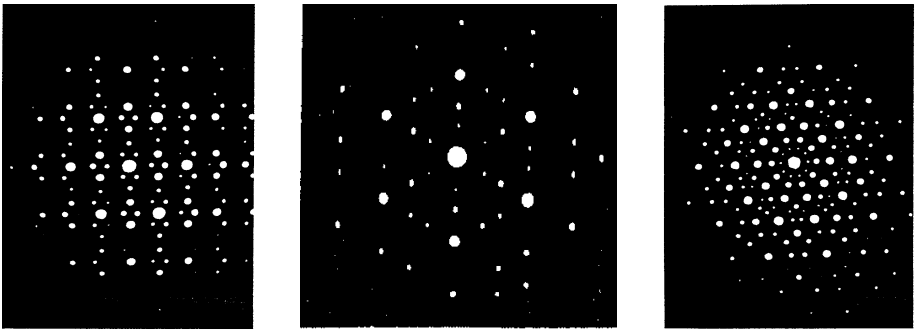


Fig. 10.6. Two-, three-, and fivefold symmetry in the electron diffraction patterns of quasicrystalline $Al_{86}Mn_{14}$ [D. Shechtman et al., *Phys. Rev. Lett.* **53**, 1951 (1984)]

Soon a broad class of materials that exhibit similar features was discovered: their diffraction patterns show symmetries that cannot be interpreted within the framework of traditional crystallography. In addition to the abovementioned $Al_{86}Mn_{14}$, icosahedral symmetry is observed in $Al_{86}Fe_{14}$, $Al_{85}Cr_{15}$, and $Al_{65}Cu_{20}M_{15}$ (where M stands for Mn, Fe, Cr, V, Ru, or Os). The appearance of icosahedral regions in quenched transition-metal alloys is in fact not surprising. For spherical atoms in a crystalline environment the closest packing is known to be offered by fcc and hcp structures of coordination twelve. For transition metals, where d orbitals play an important role, the energetically most favorable arrangement in a cluster of 13 atoms does not correspond to the local environment in fcc or hcp structures (cuboctahedra or anticuboctahedra): here the twelve nearest neighbors are arranged icosahedrally around the thirteenth, as shown in Fig. 7.9. In the new class of materials the building blocks are such icosahedral units, however their positions show no long-range order. Nevertheless some kind of long-range order, namely bond-orientational order is preserved. Although discrete translational symmetry is broken, the structure is quasiperiodic under translations. Such materials are called *quasicrystals*, abbreviated from quasiperiodic crystals.

In addition to icosahedral symmetry, other symmetries ruled out in conventional crystallography may also appear in quasicrystals. Eight-, ten- and twelvefold symmetries are observed in $Mn_{82}Si_{15}Al_3$, $Al_{65}Cu_{20}Mn_{15}$, and $V_{15}Ni_{10}Si$, respectively. Needless to say, there are many other examples for each case. Quasicrystals can be classified according to their noncrystallographic symmetries, thus we speak of octagonal, decagonal, dodecagonal, and icosahedral quasicrystals.

The absence of long-range periodic order in atomic positions is clearly shown by the partial distribution functions in Fig. 10.7. While oscillations are damped less rapidly than in amorphous materials, the radial distribution function is nevertheless closer to that of amorphous materials than of crystals. On the other hand the diffraction pattern indicates the presence of a long-range quasiperiodic order – just like in incommensurate structures. Contrary to the latter, where only crystallographically allowed rotations are observed, quasicrystals show rotational symmetries that are ruled out in 3D crystals. In fact quasiperiodicity is precisely due to such symmetries; even the shape of quasicrystals may reflect these symmetries.

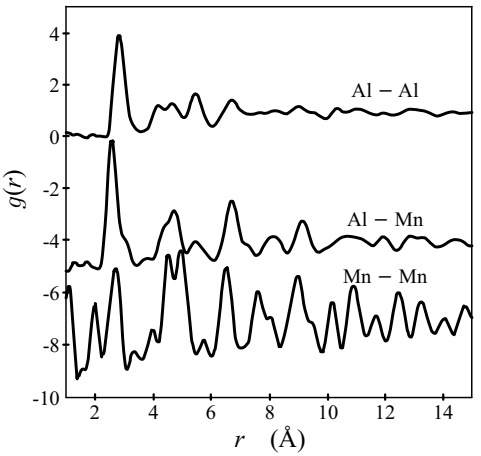


Fig. 10.7. Partial distribution functions for Al–Al, Al–Mn, and Mn–Mn in icosahedral quasicrystalline $Al_{74}Si_5Mn_{21}$, measured by M. de Boissieu et al. (*J. Phys.: Condens. Matter* **2**, 2499 (1990)) using neutron diffraction techniques. For better visibility, curves are shifted vertically

Without going into mathematical details, we shall now sketch a simple model for quasicrystalline order that also explains the appearance of relatively sharp Bragg peaks.

10.2.4 The Fibonacci Chain

We shall start the analysis of the diffraction patterns of quasicrystals with the example of a one-dimensional quasiperiodic arrangement, a chain built up of short (S) and long (L) segments according to the construction rule of the Fibonacci sequence.[2] The zeroth generation of the Fibonacci sequence consists of a single element S, and the first generation is composed of a single element L. Then, starting from the second generation, the nth generation is obtained by joining the two previous ones. The Fibonacci sequences constructed in this way using the Fibonacci rule $\Sigma_n = \Sigma_{n-1} + \Sigma_{n-2}$ are shown in Table 10.1.

Table 10.1. Construction of new generations of Fibonacci sequences by joining the two previous ones. The number of elements in the sequence is always a Fibonacci number

Fibonacci number	Fibonacci sequence
1	S
1	L
2	LS
3	LSL
5	$LSLLS$
8	$LSLLSLSL$
13	$LSLLSLSLLSLLS$
21	$LSLLSLSLLSLLSLSLLSLSL$
34	$LSLLSLSLLSLLSLSLLSLSLLSLLSLSLLSLLS$

Note that the nth generation of the sequence can be obtained from the $(n-1)$st generation by replacing each S by an L ($S \to L$), and each L by an L and S ($L \to LS$). When L and S are identified as long and short segments placed along a line, and the ratio of the lengths of the two segments is the golden mean $\tau = (1+\sqrt{5})/2 = 1.618\ldots$, then the chain generated with the above method is called the *Fibonacci chain*. The Fibonacci chain has the interesting feature that if the lengths of the segments are scaled down by τ in every iteration step, the overall length of the chain is left unchanged because of the relation

$$\frac{1}{\tau}\left(1+\frac{1}{\tau}\right) = 1\,. \tag{10.2.26}$$

Figure 10.8 shows the result of two successive iteration steps.

Now consider identical atoms placed along a line, separated by short and long distances according to the sequence in the Fibonacci chain. When diffraction is performed as a thought experiment, the resulting pattern is easily determined numerically, since according to Eq. (8.1.50) the amplitude is given

[2] Leonardo Pisano (Leonard of Pisa), commonly known as Fibonacci, 1202.

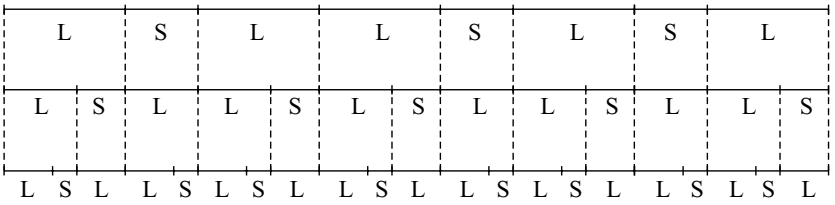


Fig. 10.8. Equal-length Fibonacci chains obtained using the rules $L \to (L+S)/\tau$ and $S \to L/\tau$

by the Fourier transform of the distribution of atoms. Because of the lack of periodicity, the Fourier components are densely distributed. Nevertheless, as shown in Fig. 10.9, sharp peaks are observed in the diffraction pattern; their positions can be specified by two indices, as expected for a one-dimensional quasiperiodic system.

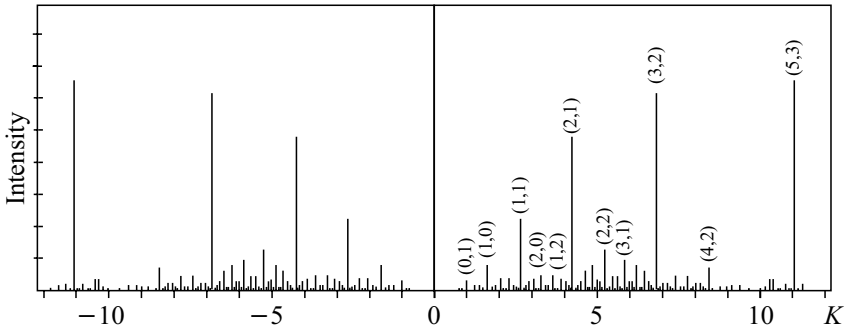


Fig. 10.9. Numerically determined diffraction pattern (Fourier spectrum) of a finite Fibonacci chain. The scale of K corresponds to $1/S$. The peaks are labeled with two integers, as explained in the text [R. D. Diehl et al., *J. Phys.: Condens. Matter* **15**, R63 (2003)]

Choosing the position of the $(1,0)$ peak as unity, the $(0,1)$ peak is found at $1/\tau$, and the peak with indices (h,k) at $h + k/\tau$. Note that there is some arbitrariness in indexing. Since the ratio of the two elementary lengths – the positions of the peaks $(1,0)$ and $(0,1)$ – is just τ, it follows from (10.2.26) that another consistent indexing is obtained by choosing the unity τ times larger or smaller. In the latter case the new indices h' and k' are related to the old ones by $h' = h + k$ and $k' = h$. This indicates the absence of a natural length scale in the Fibonacci chain – and quasicrystals in general. Note that whichever indexing is chosen, the intensity will be large for those peaks in which the two indices are subsequent Fibonacci numbers.

To demonstrate this, we shall calculate analytically the Fourier components that correspond to the quasiperiodic spectrum, and from them the

diffraction pattern. To this end we shall make use of another construction of the Fibonacci sequence. Consider the sequence of numbers $\lfloor m/\tau \rfloor$, where $\lfloor x \rfloor$ ("floor x") denotes the integer part of x, i.e., the largest integer less than or equal to x. Starting with $m = 1$, we have

$$\lfloor m/\tau \rfloor = 0, 1, 1, 2, 3, 3, 4, 4, 5, 6, 6, 7, 8, 8, 9, 9, 10, \ldots .$$

Now taking the difference between adjacent numbers,

$$\left\lfloor \frac{m+1}{\tau} \right\rfloor - \left\lfloor \frac{m}{\tau} \right\rfloor = 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 0, 1, 0, 1, \ldots . \tag{10.2.27}$$

When each 1 is replaced by L and each 0 by S, the Fibonacci chain is recovered. Thus a mathematical formula can be given for the alternation of long and short segments. The mth place is occupied by a long (L) segment if

$$\left\lfloor \frac{m+1}{\tau} \right\rfloor = \left\lfloor \frac{m}{\tau} \right\rfloor + 1 . \tag{10.2.28}$$

On the other hand, if

$$\left\lfloor \frac{m+1}{\tau} \right\rfloor = \left\lfloor \frac{m}{\tau} \right\rfloor , \tag{10.2.29}$$

then there is a short (S) segment in the mth place. Consequently, expressed in units of the length of the shorter segment, the distance of the mth point of the Fibonacci chain from the origin is

$$x_m = S \left\{ m + (\tau - 1) \left\lfloor \frac{m+1}{\tau} \right\rfloor \right\} = S \left\{ m + \frac{1}{\tau} \left\lfloor \frac{m+1}{\tau} \right\rfloor \right\} . \tag{10.2.30}$$

Using the Heaviside step function this can be alternatively written as

$$x_m = S \Big\{ m + \sum_n (\tau - 1) n \theta\big(n + 1 - (m+1)/\tau\big) \theta\big((m+1)/\tau - n\big) \Big\} . \tag{10.2.31}$$

This result is simply illustrated, and the illustration is easily generalized to the description of quasicrystals in higher dimensions. Consider a two-dimensional square lattice of lattice constant a, and mark the points (m, n) for which $n = \lfloor (m+1)/\tau \rfloor$. Then these points are projected onto the straight line of slope $\tan \phi = 1/\tau$. This is shown in Figure 10.10. It is clear from the construction that the distance of the mth point from the origin is

$$x_m = a\, m \, \cos \phi + a \, \lfloor (m+1)/\tau \rfloor \, \sin \phi . \tag{10.2.32}$$

When the lattice constant is chosen as $a = S/\cos \phi$, these distances are the same as in (10.2.30), so this construction generates a Fibonacci chain, too.

When adjacent points (m, n) in the figure are connected, long (short) segments of the Fibonacci chain are seen to be the projections of diagonal (horizontal) edges of a square cell in the lattice. Note that the selected lattice

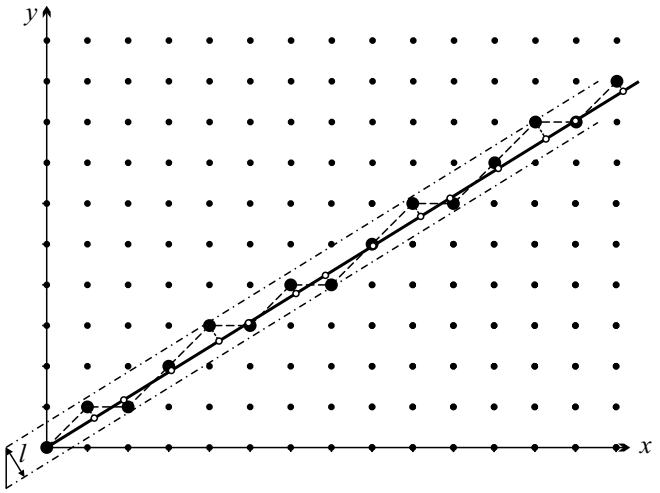


Fig. 10.10. Construction of the Fibonacci chain via the projection of selected points of a two-dimensional square lattice

points are inside a strip of width $l = a\cos\phi$ around the line of slope $1/\tau$, and the above-defined trajectory passes through all lattice points within the strip. This implies that the quasiperiodic Fibonacci chain can also be constructed by projecting the points inside a finite strip of the square lattice on a straight line with an irrational slope of $1/\tau$.

Yet another, frequently used representation of the Fibonacci chain is obtained by adding to the previously selected set of lattice points the corner points below the long diagonal sections – i.e., for every m those points (m, n) are chosen for which $n = \lfloor m/\tau \rfloor$ or $n = \lfloor (m+1)/\tau \rfloor$, and then these points, which form a quasiperiodic staircase, are projected on the line with slope $\tan\phi = 1/\tau$, as shown in Fig. 10.11. Compared to the previous staircase that was obtained by connecting the points (m, n), horizontal sections are preserved, while each diagonal section is replaced by a horizontal plus a vertical one. It follows from the scaling property of the Fibonacci chain that the projected points once again form a Fibonacci chain – however, now the projections of horizontal edges are long segments and those of vertical edges are short segments. The length of the latter is, of course, $1/\tau$ times the former, that is, $S = a\sin\phi$. This construction corresponds to the following algorithm. Elements of the Fibonacci chain are chosen in succession; for long elements (L) a step is made to the right on the square lattice, for short ones (S) a step is made upward. The staircase traced out in the two-dimensional lattice lies in a strip whose slope is $\tan\phi = 1/\tau$ and whose width l is the projection of the unit square on the direction perpendicular to the strip, i.e.,

$$l = a(\cos\phi + \sin\phi) = S(\tau + 1) = S\tau^2 \,. \tag{10.2.33}$$

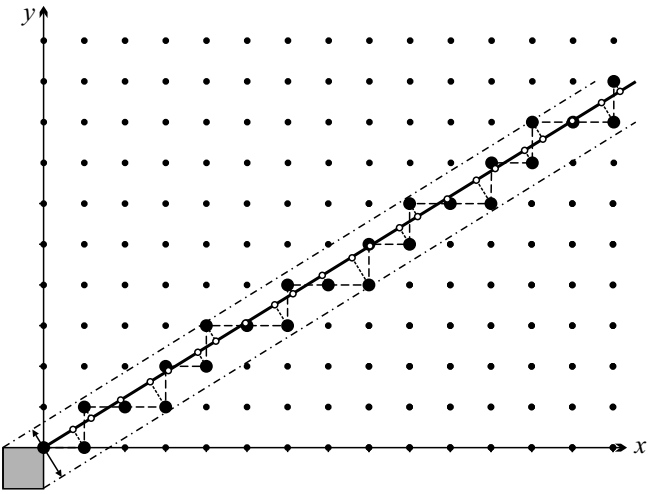


Fig. 10.11. Another construction of the Fibonacci chain via the projection of selected points of a two-dimensional square lattice

To evaluate the amplitude of the scattered beam consider the diffraction by a system that contains the points of a planar square lattice within a strip of finite width. For simplicity, we shall present the calculation in the model that represents the Fibonacci chain by the projections of points $(m, \lfloor (m+1)/\tau \rfloor)$. Using formula (10.2.31) for the coordinates,

$$\sum_m \mathrm{e}^{-\mathrm{i}Kx_m} = \sum_{m,n} \mathrm{e}^{-\mathrm{i}KS(m+n/\tau)} \theta\big(n+1-(m+1)/\tau\big)\theta\big((m+1)/\tau - n\big)\,. \tag{10.2.34}$$

When the two-dimensional vector $\boldsymbol{K} \equiv (K_x, K_y) = (K, K/\tau)$ is introduced, the previous expression can be written as

$$\begin{aligned}\sum_m \mathrm{e}^{-\mathrm{i}Kx_m} = \int \mathrm{d}\boldsymbol{r}\, \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}} \sum_{m,n} \delta(x-Sm)\delta(y-Sn) \\ \times\theta[y/S+1-(x/S+1)/\tau]\theta[(x/S+1)/\tau - y/S]\,.\end{aligned} \tag{10.2.35}$$

Next, the function to be Fourier transformed is considered as a product of two terms; the first contains a summation over all points of the square lattice, and the second is a product of step functions projecting them to the narrow strip. Their respective Fourier transforms are easily determined,

$$\begin{aligned}\int \mathrm{d}\boldsymbol{r}\, \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}} \sum_{m,n} \delta(x-Sm)\delta(y-Sn) \\ = \frac{N}{V}(2\pi)^2 \sum_{h,k} \delta(K_x - h2\pi/S)\delta(K_y - k2\pi/S)\,,\end{aligned} \tag{10.2.36}$$

and

$$\begin{aligned}\int \mathrm{d}\boldsymbol{r}\, \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}}\theta[y/S+1-(x/S+1)/\tau]\theta[(x/S+1)/\tau - y/S] \\ = \int \mathrm{d}x \int_{(x-S/\tau)/\tau}^{(x+S)/\tau} \mathrm{d}y\, \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}} \\ = \int \mathrm{d}x\, \mathrm{e}^{-\mathrm{i}K_x x} \frac{\mathrm{i}}{K_y}\left[\mathrm{e}^{-\mathrm{i}K_y(x+S)/\tau} - \mathrm{e}^{-\mathrm{i}K_y(x-S/\tau)/\tau}\right].\end{aligned} \tag{10.2.37}$$

Since the Fourier transform of the product is the convolution of the Fourier transforms of its factors,

$$\begin{aligned}\sum_m \mathrm{e}^{-\mathrm{i}Kx_m} = (2\pi)^2 \frac{N}{V} \sum_{h,k} \int \mathrm{d}x \iint \mathrm{d}q_x \mathrm{d}q_y\, \delta(K_x - q_x - h2\pi/S) \\ \times\delta(K_y - q_y - k2\pi/S)\mathrm{e}^{-\mathrm{i}q_x x} \frac{\mathrm{i}}{q_y}\left[\mathrm{e}^{-\mathrm{i}q_y(x+S)/\tau} - \mathrm{e}^{-\mathrm{i}q_y(x-S/\tau)/\tau}\right].\end{aligned} \tag{10.2.38}$$

Integration yields a factor

$$\delta[K(1+1/\tau^2) - (2\pi/S)(h+k/\tau)]\,, \tag{10.2.39}$$

indicating that diffraction peaks appear for those values of K that satisfy

$$K = K_{hk} = \frac{2\pi}{S}\frac{\tau^2}{1+\tau^2}(h+k/\tau) = \frac{2\pi}{a}\frac{1}{\sqrt{1+\tau^2}}(h\tau + k)\,, \tag{10.2.40}$$

where h and k are integers. Writing the structure amplitude as

$$F_K = \sum_{hk} F_{hk}\delta(K - K_{hk})\,, \tag{10.2.41}$$

the Fourier coefficient is found to be

$$\begin{aligned}F_{hk} &= \lim_{N\to\infty} \frac{1}{N}\sum_n \mathrm{e}^{-\mathrm{i}Kx_n} \\ &= \frac{\sin\left[\dfrac{\pi\tau}{1+\tau^2}(\tau k - h)\right]}{\dfrac{\pi\tau}{1+\tau^2}(\tau k - h)} \exp\left[\mathrm{i}\pi\frac{\tau-2}{\tau+2}(\tau k - h)\right].\end{aligned} \tag{10.2.42}$$

Thus diffraction peaks are indeed specified by two integer labels. The intensity of the diffraction peak of indices hk is found to be proportional to

$$\frac{\sin^2\left[\dfrac{\pi\tau}{1+\tau^2}(-h+\tau k)\right]}{\left[\dfrac{\pi\tau}{1+\tau^2}(-h+\tau k)\right]^2}. \tag{10.2.43}$$

The indices h and k run over all (positive and negative) integers, so, due to the irrationality of τ, the allowed values of K make up a quasicontinuous spectrum. However – as already mentioned in relation to Fig. 10.9 – the intensity is large only at peaks for which the ratio of the indices is close to τ – that is, for which the two integers are subsequent Fibonacci numbers.

To provide an intuitively clear interpretation of this result, we shall return to the representation of the Fibonacci chain by the projection of the points in a strip of a square lattice. The reciprocal lattice of the square lattice is made up by the points $\boldsymbol{K} = \frac{2\pi}{a}(h,k)$. In the coordinate system rotated by an angle ϕ these points are given as

$$\begin{aligned} K_{hk\parallel} &= \frac{2\pi}{a}(h\cos\phi + k\sin\phi) = \frac{2\pi\tau}{a\sqrt{1+\tau^2}}(h + k/\tau)\,, \\ K_{hk\perp} &= \frac{2\pi}{a}(-h\sin\phi + k\cos\phi) = \frac{2\pi}{a\sqrt{1+\tau^2}}(-h + \tau k)\,. \end{aligned} \tag{10.2.44}$$

Then, apart from the phase factors, the expression for the structure amplitude is written as

$$F_K = \sum_{h,k} \frac{\sin[K_{hk\perp}a(\cos\phi)/2]}{K_{hk\perp}a(\cos\phi)/2}\delta(K - K_{hk\parallel})\,. \tag{10.2.45}$$

This shows that the scattered intensity receives contributions from each point $\boldsymbol{K}$ of the reciprocal lattice. The scattering "vector" K, which determines the position of the diffraction peak associated with the reciprocal-lattice point $\boldsymbol{K}$ is the projection $K_{hk\parallel}$ of $\boldsymbol{K}$ on a line of irrational slope, while the amplitude of the diffraction peak is determined by its projection along the perpendicular direction. This amplitude is small when the projection is large, i.e., when $\boldsymbol{K}$ is not close to the specified line. To understand this result recall that the Fibonacci chain was constructed using the points in a narrow strip (of width $l = a\cos\phi$) in the direct lattice, therefore diffraction by the Fibonacci chain can be interpreted as scattering by a specially oriented narrow strip of a square lattice. In contrast to the usual, rather sharp diffraction peaks that appear for macroscopic samples, the peaks now have a finite width in the direction perpendicular to the strip. The diffraction pattern is shown in Fig. 10.12. Diffraction peaks far from the line $K_\perp = 0$ hardly contribute. This leads to a relatively sparse distribution of sharp peaks.

10.2.5 Penrose Tiling of the Plane

The two-dimensional generalization of the quasiperiodic Fibonacci chain is the quasiperiodic covering (tiling) of the plane with two objects. As demonstrated by R. Penrose in 1974, this is possible with two rhombi of equal sides. The angles of the thinner rhombus are 36° and 144°, while those of the thicker are 72° and 108°. These figures are called *Penrose tiles*. To rule out periodic tilings, sides are marked with single or double arrows, and the common sides

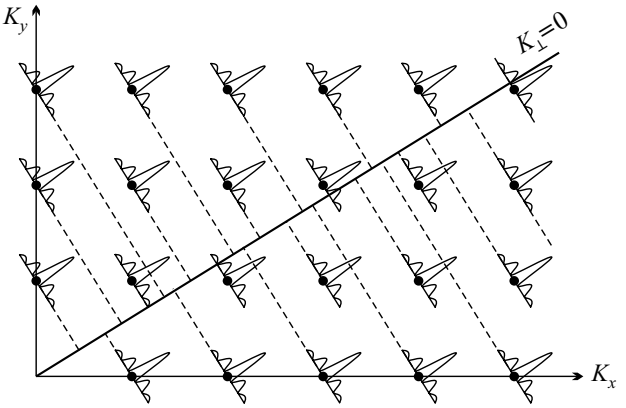


Fig. 10.12. Diffraction pattern of a finite strip of a square lattice and its section along the line $K_\perp = 0$

of neighboring tiles are required to have the same number of arrows in the same direction. Figure 10.13 shows the elementary rhombi and some simple allowed matches.

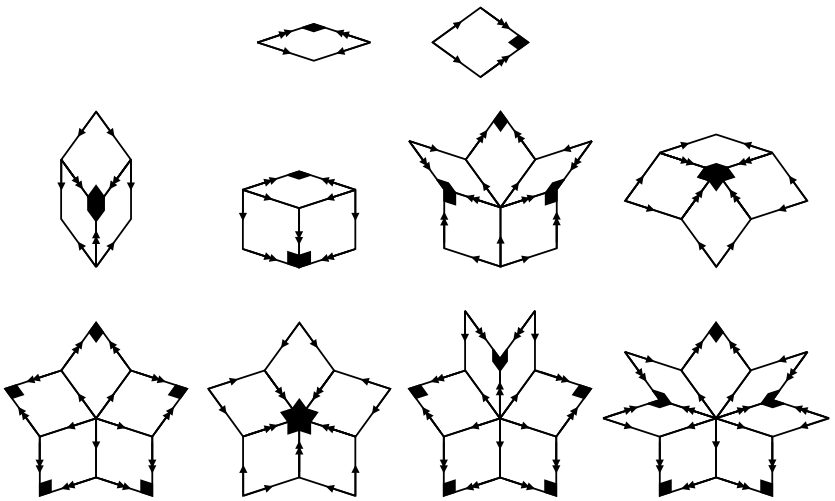


Fig. 10.13. The 36° and 72° rhombi used in Penrose tiling and some simple matches

The rule for matching tiles is often expressed in an alternative way: each tile has a marked vertex – the one where the two sides with double arrows join –, and tiles can be matched only if the joining vertices are either all marked or all unmarked. Figure 10.14 shows a possible tiling of the plane that

respects matching rules. For clarity, thin tiles are shaded gray. The presence of local fivefold symmetry in many places is obvious at first sight. In spite of the lack of strict long-range periodicity, the tile edges (or bonds between atoms at the vertices) are all along five specific directions, thus the pattern shows long-range directional order.

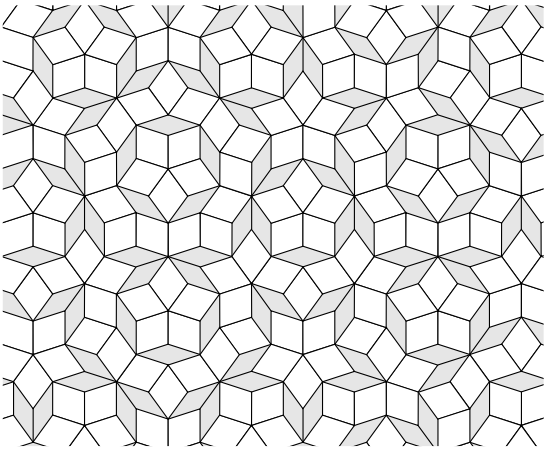


Fig. 10.14. Penrose tiling of the plane using 36° and 72° rhombi

The position vectors of the vertices of the rhombi can all be written as

$$\boldsymbol{r}_n = \sum_{i=1}^{5} n_i \boldsymbol{e}_i \,, \tag{10.2.46}$$

where the n_i are integers and the

$$\boldsymbol{e}_i = a\left(\cos\frac{2\pi(i-1)}{5}, \sin\frac{2\pi(i-1)}{5}\right), \qquad i = 1, 2, \ldots, 5 \tag{10.2.47}$$

are the vectors drawn from the center of a circle of radius a to the vertices of an inscribed regular pentagon. Two neighboring vectors $\boldsymbol{e}_i$ plus their resultant and the origin make up a thick rhombus, while second-neighbors give a thin rhombus. The five vectors are obviously not linearly independent, e.g.,

$$\boldsymbol{e}_1 = \tau(\boldsymbol{e}_2 + \boldsymbol{e}_5)\,. \tag{10.2.48}$$

If the n_i in (10.2.46) could take any integer values, points $\boldsymbol{r}_n$ would fill the plane densely because of the irrationality of τ. To obtain a Penrose tiling, only certain integers are allowed, ensuring finite distances between lattice points.

To establish the selection rule for the integers note that if the primitive vectors of a five-dimensional hypercubic lattice are projected on a particular

plane that is perpendicular to the space diagonal [11111] and spanned by two vectors with irrational components,

$$\begin{aligned} \boldsymbol{a} &= (1, \cos 2\pi/5, \cos 4\pi/5, \cos 6\pi/5, \cos 8\pi/5)\,, \\ \boldsymbol{b} &= (0, \sin 2\pi/5, \sin 4\pi/5, \sin 6\pi/5, \sin 8\pi/5)\,, \end{aligned} \tag{10.2.49}$$

the vectors (10.2.47) are recovered. Then, similarly to the Fibonacci chain, the Penrose tiling of the plane with the two rhombi can be obtained by translating the five-dimensional unit cube along this plane and projecting the lattice points within the covered finite strip onto the plane. The obtained pattern does not contain every integral linear combination of the vectors $\boldsymbol{e}_i$ – only those that make up the Penrose tiling.

The projections $\boldsymbol{q}_i$ of the reciprocal-lattice vectors[3] onto the plane are used to specify the vectors for which the structure amplitude is finite. Since the vectors $\boldsymbol{q}_i$ have irrational components, the Penrose tiling is quasiperiodic. The diffraction peaks appear at

$$\boldsymbol{K} = \sum_{i=1}^{5} h_i \boldsymbol{q}_i\,; \tag{10.2.50}$$

they are indexed by five parameters. Moreover, since the arrangement of the vectors $\boldsymbol{q}_i$ is similar to that of the $\boldsymbol{e}_i$, the diffraction pattern of the Penrose tiling shown in Fig. 10.15 exhibits five- and tenfold symmetries.

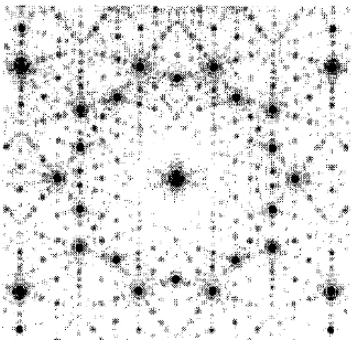


Fig. 10.15. Calculated diffraction pattern of a two-dimensional quasicrystalline arrangement of points that corresponds to a Penrose tiling [M. Senechal, *Quasicrystals and Geometry*, Cambridge University Press (1995)]

Note that matching rules that permit only quasiperiodic tilings of the plane can be established for other choices of tiles as well, e.g., a lozenge and a square (i.e., of 45° and 90° rhombi), or three rhombi with acute angles 30°, 60°, and 90° with equal side length. These choices lead to two-dimensional quasicrystals that exhibit four- and sixfold symmetries.

[3] The reciprocal lattice of a five-dimensional hypercubic lattice is also hypercubic.

10.2.6 Three-Dimensional Quasicrystals

In three dimensions quasiperiodic filling of space with simple objects is much more difficult to illustrate, therefore we shall adopt the opposite approach. First we shall investigate what symmetries may be exhibited by the structures obtained via the projection of a part of a higher-dimensional lattice onto three dimensions, and only then shall we turn to the spatial arrangement of atoms in three-dimensional quasicrystals.

Quasicrystals with icosahedral symmetry make up the most characteristic class of three dimensional quasicrystals. Their diffraction patterns display two-, three-, as well as fivefold rotation axes. The icosahedral point groups I (235) and I_h ($m\bar{3}\bar{5}$) are known to be incompatible with three-dimensional spatial periodicity. However in six-dimensional space the six fivefold axes of the icosahedral point group appear naturally. As a generalization of the foregoing, icosahedral quasicrystals may therefore be regarded as the three-dimensional projections of a part of a six-dimensional cubic lattice.

For a precise mathematical formulation, consider the projection of a six-dimensional hypercubic lattice on a particular three-dimensional subspace that is perpendicular to the direction [111111] and spanned by the vectors

$$\begin{aligned}
\boldsymbol{a} &= \frac{2}{\sqrt{5}}(1, \cos 2\pi/5, \cos 4\pi/5, \cos 6\pi/5, \cos 8\pi/5, 0)\,, \\
\boldsymbol{b} &= \frac{2}{\sqrt{5}}(0, \sin 2\pi/5, \sin 4\pi/5, \sin 6\pi/5, \sin 8\pi/5, 0)\,, \\
\boldsymbol{c} &= \frac{1}{\sqrt{5}}(-1, -1, -1, -1, -1, \sqrt{5})\,.
\end{aligned} \tag{10.2.51}$$

The projections of the primitive vectors $a(100000)$, $a(010000)$, ... of the hypercubic lattice are the vectors

$$\boldsymbol{e}_i = \frac{2a}{\sqrt{5}}\left(\cos\frac{2\pi(i-1)}{5},\ \sin\frac{2\pi(i-1)}{5},\ -\frac{1}{2}\right), \quad i = 1, 2, \ldots, 5 \tag{10.2.52}$$

and $\boldsymbol{e}_6 = a(0, 0, 1)$. It is readily seen that these projected vectors have the same magnitude, too. For an even clearer manifestation of the symmetries the vectors are rotated around the y-axis through $\phi = \arctan(1/\tau)$. When this new set of basis vectors,

$$\begin{aligned}
\boldsymbol{e}_1 &= \frac{a}{\sqrt{1+\tau^2}}(1, 0, -\tau)\,, & \boldsymbol{e}_2 &= \frac{a}{\sqrt{1+\tau^2}}(0, \tau, -1)\,, \\
\boldsymbol{e}_3 &= \frac{a}{\sqrt{1+\tau^2}}(-\tau, 1, 0)\,, & \boldsymbol{e}_4 &= \frac{a}{\sqrt{1+\tau^2}}(-\tau, -1, 0)\,, \\
\boldsymbol{e}_5 &= \frac{a}{\sqrt{1+\tau^2}}(0, -\tau, -1)\,, & \boldsymbol{e}_6 &= \frac{a}{\sqrt{1+\tau^2}}(1, 0, \tau)
\end{aligned} \tag{10.2.53}$$

and their opposites are all drawn from a common origin, their tips are seen to point into the vertices of the regular icosahedron inscribed in a sphere of radius a (see Fig. 5.12).

The primitive vectors of the reciprocal lattice of a six-dimensional hypercubic lattice generate another 6D hypercubic lattice. Projection on three-dimensional space and then rotation leads to the vectors

$$\begin{aligned}
\boldsymbol{q}_1 &= \frac{2\pi}{a\sqrt{1+\tau^2}}(1,0,-\tau)\,, \qquad \boldsymbol{q}_2 = \frac{2\pi}{a\sqrt{1+\tau^2}}(0,\tau,-1)\,,\\
\boldsymbol{q}_3 &= \frac{2\pi}{a\sqrt{1+\tau^2}}(-\tau,1,0)\,, \qquad \boldsymbol{q}_4 = \frac{2\pi}{a\sqrt{1+\tau^2}}(-\tau,-1,0)\,,\\
\boldsymbol{q}_5 &= \frac{2\pi}{a\sqrt{1+\tau^2}}(0,-\tau,-1)\,, \qquad \boldsymbol{q}_6 = \frac{2\pi}{a\sqrt{1+\tau^2}}(1,0,\tau)\,.
\end{aligned} \tag{10.2.54}$$

Diffraction peaks therefore appear at those values of $\boldsymbol{K}$ that can be written as

$$\boldsymbol{K} = \sum_{i=1}^{6} h_i \boldsymbol{q}_i\,, \tag{10.2.55}$$

where the h_i are integers. A simple rearrangement of the terms leads to the form

$$\boldsymbol{K} = \frac{2\pi}{a\sqrt{1+\tau^2}}(h\tau + h', k\tau + k', l\tau + l')\,. \tag{10.2.56}$$

The distribution of the diffraction peaks of an icosahedral quasicrystal is such that along each of the three directions the observed diffraction pattern is similar to that of a Fibonacci chain. Consequently, diffraction peaks are specified by six indices. In principle, arbitrarily close-lying Bragg peaks should also be observed, however only a few of the peaks are sufficiently intense. This is clearly seen on the powder diffraction pattern shown in Fig. 10.16. In comparison, the diffraction pattern obtained after annealing is shown in the bottom part of the figure. Diffraction peaks in the annealed sample correspond to scattering by orthorhombic Al_6Mn, while the Bragg peaks of the quenched sample may only be interpreted in terms of a quasicrystalline structure, and should be indexed by six parameters.

It is much more difficult to visualize the atomic arrangement in a three-dimensional quasicrystal than in the one- and two-dimensional cases. The 3D generalization of planar Penrose tiles is a pair of rhombohedra, spanned by the vectors given in (10.2.52). From the arrangement of the vectors it is easy to show that one type is defined by three vectors pointing to three adjacent vertices of the icosahedron, e.g., $-\boldsymbol{e}_1$, $-\boldsymbol{e}_2$, and $\boldsymbol{e}_6$, while the other by three vectors pointing to three nonadjacent vertices, e.g., $\boldsymbol{e}_1$, $\boldsymbol{e}_3$, and $\boldsymbol{e}_6$. These rhombohedra are shown in Fig. 10.17. The volume ratio of the two rhombohedra is precisely τ, so they are often referred to as "thin" and "thick". When matching conditions are suitably chosen, only quasiperiodic filling of the space is possible with them.

Figure 10.18 shows the diffraction patterns calculated from the Fourier spectra of quasicrystals obtained by filling the space with such rhombohedra. The three patterns correspond to different orientations.

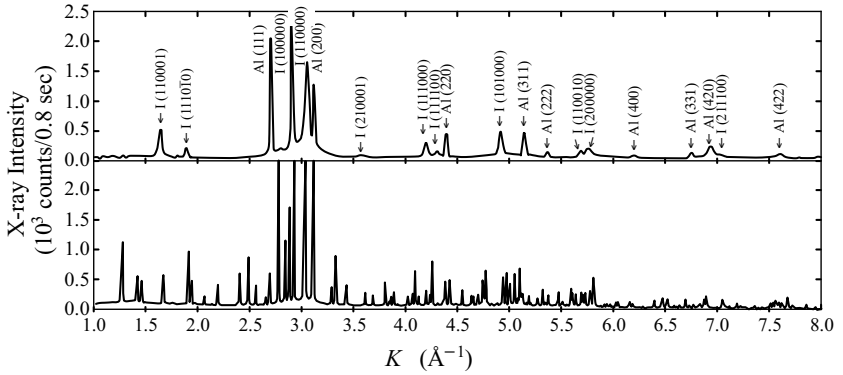


Fig. 10.16. High-resolution X-ray diffraction pattern of quenched (quasicrystalline) and annealed (crystalline) Al-Mn powder. Labels are Al and icosahedral Miller indices [P. A. Bancel et al., *Phys. Rev. Lett.* **54**, 2422 (1985)]

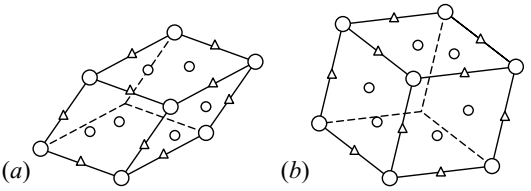


Fig. 10.17. (*a*) Thin and (*b*) thick rhombohedra used for the Penrose tiling of three-dimensional space. In quasicrystals atoms are located at marked positions, in vertices, on edges, or on face diagonals

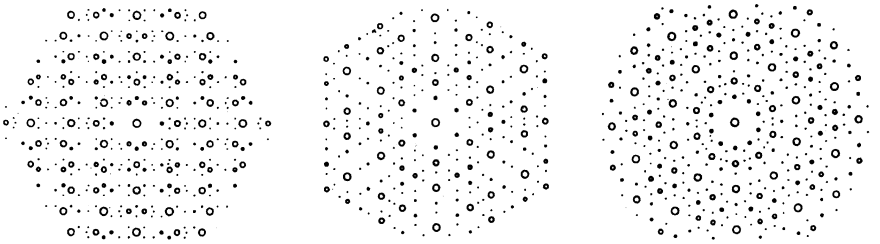


Fig. 10.18. Calculated Laue diffraction patterns of a quasicrystal constructed with thin and thick rhombohedra. The three patterns correspond to different orientations [A. Katz and M. Duneau, *J. Physique* **47**, 181 (1986)]

The similarity with the experimental results shown in Fig. 10.6 is striking. To achieve perfect agreement, the rhombohedra have to be suitably decorated with atoms, in line with the chemical composition of the quasicrystal. Figure 10.17 also shows the sites where atoms are expected to be located. In quasicrystalline $Al_{86}Mn_{14}$ manganese atoms occupy the vertices and alu-

minum atoms all other marked sites – nevertheless occupation should always be understood in an average sense.

Just as real crystals are never perfectly periodic because of the large number of defects they contain, orientational order always extends only over finite regions in quasicrystals as well. Defects and disordered regions inevitably appear in them.

Further Reading

1. N. E. Cusack, *The Physics of Structurally Disordered Matter: An Introduction*, Graduate Student Series in Physics, Adam Hilger, Bristol (1987).
2. *Introduction to Quasicrystals*, Edited by M. V. Jarić, Academic Press, Inc., Boston (1988).
3. C. Janot, *Quasicrystals, A Primer*, Second Edition, Clarendon Press, Oxford (1997).
4. G. Venkataraman, D. Sahoo, and V. Balakrishnan, *Beyond the Crystalline State: An Emerging Perspective*, Springer Series in Solid-State Sciences, Vol. 84., Springer-Verlag Berlin (1989).
5. Y. Waseda, *The Structure of Non-Crystalline Materials: Liquids and Amorphous Solids*, McGraw-Hill, New York (1980).
6. R. Zallen, *The Physics of Amorphous Solids*, John Wiley & Sons, New York (1998).
7. J. M. Ziman, *Models of Disorder: The Theoretical Physics of Homogeneously Disordered Systems*, Cambridge University Press, Cambridge (1979).

11

Dynamics of Crystal Lattices

The discussion of the structure of solids was based on the assumption that atoms or ions are rigidly fixed at their equilibrium positions. Within the framework of classical physics this may be valid in the ground state of the crystal. However, at finite temperatures atoms are displaced by thermal motion from their equilibrium positions. Without taking this into consideration, certain mechanical, elastic, and thermal properties of solids could not be properly accounted for. For example, barring the regime of low temperatures, the calculated specific heat is found to be much smaller than the observed value when the vibration of ions is ignored and only the contribution of electrons is retained. Similarly, neither thermal expansion nor melting can be explained in a model that assumes the crystal structure to be rigid – not to mention mechanical properties, which do not lend themselves to interpretation in terms of a rigid lattice. Besides, lattice dynamics modifies the electronic states as well, as we shall see in the Volume 2.

Making use of the vacancies and interstitials in the lattice, atoms may move far from their initial position. In this chapter we shall ignore this diffusion and assume that atoms stay close to their equilibrium positions (which correspond to a crystalline order), and perform small-amplitude oscillations about it. Here we shall treat the problem classically; quantum effects will be taken into account in the next chapter.

11.1 The Harmonic Approximation

In principle, the motion of the atoms in a lattice can be determined unambiguously if the potential felt by them is known. In insulators only pairwise interactions among the atoms (interatomic pair potentials) need to be considered. In metals the full interaction is the sum of the direct Coulomb potentials due to other ions and the potential due to highly mobile delocalized electrons that are not bound to any ion. The interaction Hamiltonian therefore depends on the coordinates of ions and mobile electrons alike. We shall demonstrate

in Chapter 23 that, owing to the large mass difference between ions and electrons, electrons follow ions practically without any delay, while ions feel only an average of the potential due to electrons. In what follows we shall therefore assume that the potential U felt by the ions depends only on the ionic coordinates, just like in insulators.

For completeness we shall consider a lattice with a basis of p atoms at positions $\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_p$ relative to the primitive cell; the lattice vectors of the primitive cells are denoted by $\boldsymbol{R}_m$. The equilibrium positions of the atoms are

$$\boldsymbol{r}^0(m,\mu) = \boldsymbol{R}_m + \boldsymbol{r}_\mu\,, \qquad \mu = 1, 2, \ldots, p. \tag{11.1.1}$$

When atoms perform thermal motion about these positions, their instantaneous coordinates at time t are given by

$$\boldsymbol{r}(m,\mu,t) = \boldsymbol{R}_m + \boldsymbol{r}_\mu + \boldsymbol{u}(m,\mu,t)\,, \tag{11.1.2}$$

where $\boldsymbol{u}(m,\mu,t)$ is the instantaneous displacement of the μth atom in the primitive cell at $\boldsymbol{R}_m$ from its equilibrium position. From now on, we shall usually suppress the time variable.

Obviously, the potential U is a function of the true atomic positions, and therefore of the displacements $\boldsymbol{u}(m,\mu)$. To obtain an easily tractable problem, further assumptions have to be made about the interactions between ions.

11.1.1 Second-Order Expansion of the Potential

When the diffusion of atoms is ignored, their displacement from equilibrium can be assumed to be small as long as the temperature is low compared to the melting point. The potential is then expanded in powers of the displacements $\boldsymbol{u}(m,\mu)$ as

$$\begin{aligned} U(\{\boldsymbol{r}(m,\mu)\}) = U_0 &+ \sum_{m,\mu,\alpha} \Phi^{\mu}_{\alpha}(m) u_\alpha(m,\mu) \\ &+ \tfrac{1}{2} \sum_{\substack{m,\mu,\alpha \\ n,\nu,\beta}} \Phi^{\mu\nu}_{\alpha\beta}(m,n) u_\alpha(m,\mu) u_\beta(n,\nu) + \ldots\,, \end{aligned} \tag{11.1.3}$$

where

$$\begin{aligned} \Phi^{\mu}_{\alpha}(m) &= \frac{\partial U(\{\boldsymbol{r}(m,\mu)\})}{\partial u_\alpha(m,\mu)}\,, \\ \Phi^{\mu\nu}_{\alpha\beta}(m,n) &= \frac{\partial^2 U(\{\boldsymbol{r}(m,\mu)\})}{\partial u_\alpha(m,\mu)\partial u_\beta(n,\nu)}\,. \end{aligned} \tag{11.1.4}$$

The Greek indices α, β label Cartesian coordinates.

The ground-state energy U_0 – which corresponds to the configuration in which every atom is in its equilibrium position – is essential for the calculation of the total energy of the solid but unimportant for the determination of the

frequency spectrum of vibrations, which is why it will be ignored below. Since the potential attains its minimum at the equilibrium positions of the atoms, the coefficients $\Phi^{\mu}_{\alpha}(m)$ of the terms linear in the displacement vanish. The approximation in which only second-order terms are kept is called the *harmonic approximation*. Throughout this chapter we shall use this approximation, as a relatively good description of the thermal properties of solids can be based on it. Phenomena that cannot be interpreted in this framework are presented at the end of the next chapter. In this context the role of higher-order terms in the expansion is also discussed there.

The ab initio determination of the coefficients Φ in the expansion is very difficult, and requires highly time-consuming numerical calculations. Therefore these coefficients are often considered as phenomenological parameters whose values are obtained from fits to experimental data. We shall also adopt this approach. Nevertheless there exist completely general relations among the coefficients through which the number of fitting parameters can be reduced substantially. One such relation is the consequence of the definition of the coefficients. As they appear as the second partial derivatives of the energy,

$$\Phi^{\mu\nu}_{\alpha\beta}(m,n) = \Phi^{\nu\mu}_{\beta\alpha}(n,m)\,. \tag{11.1.5}$$

Further relations are obtained from the expression of the force on individual atoms. The force on the atom labeled (m,μ) is derived from the potential as

$$\boldsymbol{F}(m,\mu) = -\frac{\partial U}{\partial \boldsymbol{u}(m,\mu)}\,. \tag{11.1.6}$$

In the harmonic approximation

$$F_{\alpha}(m,\mu) = -\sum_{n,\nu,\beta} \Phi^{\mu\nu}_{\alpha\beta}(m,n) u_{\beta}(n,\nu)\,. \tag{11.1.7}$$

Suppose that each atom is displaced by $\boldsymbol{u}^0$, which corresponds to a homogeneous translation of the crystal. Since atoms are not displaced relative to each another, the crystal remains in equilibrium, and the force on each atom is zero; hence

$$\sum_{n,\nu} \Phi^{\mu\nu}_{\alpha\beta}(m,n) = 0\,. \tag{11.1.8}$$

From (11.1.5) it follows that

$$\sum_{m,\mu} \Phi^{\mu\nu}_{\alpha\beta}(m,n) = 0\,. \tag{11.1.9}$$

By making use of the two previous formulas it is straightforward to show that the second-order term in the expansion of the potential may be written in the equivalent form

$$-\tfrac{1}{4} \sum_{\substack{m,\mu,\alpha \\ n,\nu,\beta}} \Phi^{\mu\nu}_{\alpha\beta}(m,n) \left[u_{\alpha}(m,\mu) - u_{\alpha}(n,\nu)\right] \left[u_{\beta}(m,\mu) - u_{\beta}(n,\nu)\right]\,. \tag{11.1.10}$$

Yet another set of connections may be obtained by exploiting the property that a rigid rotation of the entire crystal does not generate any internal forces either. Even more important is that in specific cases knowledge of the crystal symmetries may substantially simplify calculations or the evaluation of measurements, as the same symmetries are shown by the potential U and so the coefficients $\Phi^{\mu\nu}_{\alpha\beta}$ as well. This possibility is nonetheless ignored in the general treatment.

As it will prove to be highly important for the applications below, it should be noted that in ordered crystals the coefficients $\Phi^{\mu\nu}_{\alpha\beta}(m,n)$ depend only on $\boldsymbol{R}_m - \boldsymbol{R}_n$, the difference of the lattice vectors $\boldsymbol{R}_m$ and $\boldsymbol{R}_n$ of the primitive cells. Denoting this by $m-n$,

$$\Phi^{\mu\nu}_{\alpha\beta}(m,n) = \Phi^{\mu\nu}_{\alpha\beta}(m-n)\,. \tag{11.1.11}$$

11.1.2 Expansion of the Energy for Pair Potentials

In many – but by no means all – cases it may be assumed that the interaction among ions can be written as the sum of pair potentials that depend only on the difference of the position vectors of the two ions:

$$\begin{aligned} U &= \tfrac{1}{2}\sum_{\substack{m,\mu\\ n,\nu}} U_{\text{pair}}\left(\boldsymbol{r}(m,\mu) - \boldsymbol{r}(n,\nu)\right) \\ &= \tfrac{1}{2}\sum_{\substack{m,\mu\\ n,\nu}} U_{\text{pair}}\left(\boldsymbol{R}_m + \boldsymbol{r}_\mu + \boldsymbol{u}(m,\mu) - \boldsymbol{R}_n - \boldsymbol{r}_\nu - \boldsymbol{u}(n,\nu)\right). \end{aligned} \tag{11.1.12}$$

Expansion around the equilibrium position is then performed, and once again terms are kept only up to second order:

$$\begin{aligned} U = {} & \tfrac{1}{2}\sum_{\substack{m,\mu\\ n,\nu}} U_{\text{pair}}(\boldsymbol{R}_m + \boldsymbol{r}_\mu - \boldsymbol{R}_n - \boldsymbol{r}_\nu) \\ & + \tfrac{1}{2}\sum_{\substack{m,\mu,\alpha\\ n,\nu}} \widetilde{\Phi}^{\mu\nu}_{\alpha}(m,n)\left[u_\alpha(m,\mu) - u_\alpha(n,\nu)\right] \\ & + \tfrac{1}{4}\sum_{\substack{m,\mu,\alpha\\ n,\nu,\beta}} \left[u_\alpha(m,\mu) - u_\alpha(n,\nu)\right]\widetilde{\Phi}^{\mu\nu}_{\alpha\beta}(m,n)\left[u_\beta(m,\mu) - u_\beta(n,\nu)\right], \end{aligned} \tag{11.1.13}$$

where $\widetilde{\Phi}$ may now be expressed in terms of the first and second partial derivatives of the pair potential. Using the concise notation $\boldsymbol{u} = \boldsymbol{u}(m,\mu) - \boldsymbol{u}(n,\nu)$, we have

$$\begin{aligned} \widetilde{\Phi}^{\mu\nu}_{\alpha}(m,n) &= \frac{\partial U_{\text{pair}}(m,\mu;n,\nu;\boldsymbol{u})}{\partial u_\alpha}\,, \\ \widetilde{\Phi}^{\mu\nu}_{\alpha\beta}(m,n) &= \frac{\partial^2 U_{\text{pair}}(m,\mu;n,\nu;\boldsymbol{u})}{\partial u_\alpha \partial u_\beta}\,. \end{aligned} \tag{11.1.14}$$

The term linear in the atomic displacements vanishes again, as the sum of the pair potentials has its minimum at the equilibrium distance of the ions. Comparing the second-order term with the expression in (11.1.10),

$$\Phi^{\mu\nu}_{\alpha\beta}(m,n) = -\widetilde{\Phi}^{\mu\nu}_{\alpha\beta}(m,n)\,, \tag{11.1.15}$$

unless $m = n$ and $\mu = \nu$. To determine the contribution of the term $m = n$, $\mu = \nu$, a consequence of the relation (11.1.8) has to be exploited:

$$\Phi^{\mu\nu}_{\alpha\beta}(m,n) = \delta_{mn}\delta_{\mu\nu} \sum_{n',\nu'} \widetilde{\Phi}^{\mu\nu'}_{\alpha\beta}(m,n') - \widetilde{\Phi}^{\mu\nu}_{\alpha\beta}(m,n)\,. \tag{11.1.16}$$

Forces are said to be central when the pair interaction depends only on the distance between the atoms:

$$U_{\rm pair}(m,\mu;n,\nu;\boldsymbol{u}) = U_{\rm pair}(|\boldsymbol{R}_m+\boldsymbol{r}_\mu+\boldsymbol{u}(m,\mu)-\boldsymbol{R}_n-\boldsymbol{r}_\nu-\boldsymbol{u}(n,\nu)|)\,. \tag{11.1.17}$$

Expansion of the potential about the equilibrium distance $r = |\boldsymbol{r}| = |\boldsymbol{R}_m + \boldsymbol{r}_\mu - \boldsymbol{R}_n - \boldsymbol{r}_\nu|$ gives

$$\frac{\partial U_{\rm pair}(|\boldsymbol{r}+\boldsymbol{u}|)}{\partial u_\alpha} = \frac{\partial U_{\rm pair}(|\boldsymbol{r}+\boldsymbol{u}|)}{\partial|\boldsymbol{r}+\boldsymbol{u}|}\frac{r_\alpha + u_\alpha}{r}\,, \tag{11.1.18}$$

which implies

$$\widetilde{\Phi}^{\mu\nu}_{\alpha\beta}(m-n) = \delta_{\alpha\beta}\frac{1}{r}\frac{\partial U_{\rm pair}(r)}{\partial r} + \left(\frac{\partial^2 U_{\rm pair}(r)}{\partial r^2} - \frac{1}{r}\frac{\partial U_{\rm pair}(r)}{\partial r}\right)\frac{r_\alpha r_\beta}{r^2}\,. \tag{11.1.19}$$

The first derivatives of the individual pair potentials do not generally vanish at the equilibrium position: only the full potential has its minimum there, which is why the first derivative appears in the expression for $\widetilde{\Phi}$. The previous formula clearly shows that apart from an isotropic term the direction of the displacement relative to the axis joining the atoms is important. Restoring forces act only when the relative displacement of the atoms has a nonvanishing projection along this axis. Displacements perpendicular to it give only second-order corrections to the interatomic distance, and therefore do not contribute to the energy.

11.1.3 Equations Governing Lattice Vibrations

The classical motion of atoms is most easily described using the Hamiltonian equations of classical mechanics. The Hamiltonian is specified by expressing the kinetic and potential energies in terms of the displacements $\boldsymbol{u}(m,\mu)$ and the conjugate momenta. Assuming that each primitive cell contains an atom of mass M_μ at the μth position, the kinetic energy due to atomic vibrations is

$$T_{\rm kin} = \sum_{m,\mu} \tfrac{1}{2} M_\mu \dot{\boldsymbol{u}}^2(m,\mu)\,. \tag{11.1.20}$$

The potential energy shall either be written as

$$U_{\text{harm}} = \tfrac{1}{2} \sum_{\substack{m,\mu,\alpha \\ n,\nu,\beta}} \Phi^{\mu\nu}_{\alpha\beta}(m,n) u_\alpha(m,\mu) u_\beta(n,\nu)\,, \tag{11.1.21}$$

or in terms of the pair potentials as

$$U_{\text{harm}} = \tfrac{1}{4} \sum_{\substack{m,\mu,\alpha \\ n,\nu,\beta}} [u_\alpha(m,\mu) - u_\alpha(n,\nu)]\, \widetilde{\Phi}^{\mu\nu}_{\alpha\beta}(m,n)\, [u_\beta(m,\mu) - u_\beta(n,\nu)]\,. \tag{11.1.22}$$

In classical mechanics the canonical momentum $\boldsymbol{P}$ conjugate to the displacement $\boldsymbol{u}$ is derived from the Lagrangian $\mathcal{L} = T_{\text{kin}} - U_{\text{harm}}$ as

$$\boldsymbol{P}(m,\mu) = \frac{\partial \mathcal{L}}{\partial \dot{\boldsymbol{u}}(m,\mu)}\,. \tag{11.1.23}$$

The kinetic energy then takes the form

$$T_{\text{kin}} = \sum_{m,\mu} \frac{\boldsymbol{P}^2(m,\mu)}{2M_\mu} = \sum_{m,\mu,\alpha} \frac{P^2_\alpha(m,\mu)}{2M_\mu}\,. \tag{11.1.24}$$

From the classical Hamiltonian, which is the sum of the kinetic and potential energies – both expressed in terms of the canonical variables –,

$$\mathcal{H} = T_{\text{kin}} + U_{\text{harm}}\,, \tag{11.1.25}$$

the following equations of motion are derived:

$$\begin{aligned} \dot{u}_\alpha(m,\mu) &= \frac{\partial \mathcal{H}}{\partial P_\alpha(m,\mu)} = \frac{P_\alpha(m,\mu)}{M_\mu}\,, \\ \dot{P}_\alpha(m,\mu) &= -\frac{\partial \mathcal{H}}{\partial u_\alpha(m,\mu)} = -\sum_{n,\nu,\beta} \Phi^{\mu\nu}_{\alpha\beta}(m-n) u_\beta(n,\nu)\,, \end{aligned} \tag{11.1.26}$$

or

$$\dot{P}_\alpha(m,\mu) = -\sum_{n,\nu,\beta} \widetilde{\Phi}^{\mu\nu}_{\alpha\beta}(m-n)[u_\beta(m,\mu) - u_\beta(n,\nu)]\,. \tag{11.1.27}$$

Differentiation of the equation for u leads to the well-known Newtonian equation

$$M_\mu \ddot{u}_\alpha(m,\mu) = -\sum_{n,\nu,\beta} \Phi^{\mu\nu}_{\alpha\beta}(m-n) u_\beta(n,\nu)\,, \tag{11.1.28}$$

or alternatively

$$M_\mu \ddot{u}_\alpha(m,\mu) = -\sum_{n,\nu,\beta} \widetilde{\Phi}^{\mu\nu}_{\alpha\beta}(m-n)[u_\beta(m,\mu) - u_\beta(n,\nu)]\,. \tag{11.1.29}$$

Since the potential is assumed to be harmonic, the equations are identical to those of a mechanical mass–spring system, which is why the quantities

$\Phi^{\mu\nu}_{\alpha\beta}(m-n)$ are called *spring constants* or *force constants*. The resulting many-variable system of coupled differential equations is seemingly very complex, however when plausible assumptions are made about the spring constants, it turns out to be solvable in some cases. The obtained intuitive picture facilitates the interpretation of the quantum mechanical results.

11.2 Vibrational Spectra of Simple Lattices

By considering the classical crystal as a system built up of Np mass points, the vibrations of the lattice are determined from the coupled system of equations (11.1.28) and (11.1.29) in $3Np$ variables. We shall first deal with some simple cases where calculations are straightforward. It will be assumed that atoms are located along a line, making up a chain, and their displacements are also in the same direction. In the simplest case the chain is made up of a single kind of atom, and nearest-neighbor distances are all identical in the equilibrium configuration. This model is called the *monatomic linear chain*. In a second, somewhat more complicated situation the chain is made up of two kinds of atoms of unequal mass, located at alternate positions. A similar treatment is applied in the case when the chain contains identical atoms and their equilibrium separations alternate. This system is called a *dimerized chain*. After the determination of the vibrational modes in these models we shall turn to the study of the vibrational spectra of simple cubic lattices, and then to the general discussion of classical vibrations in three-dimensional lattices.

11.2.1 Vibrations of a Monatomic Linear Chain

Consider a linear chain of lattice constant a with atoms of mass M at the lattice points. The origin of coordinates is chosen in such a way that the equilibrium position of the nth atom is na. The displacement u_n of each vibrating atom is supposed to be along the chain. The atoms, their equilibrium positions and displacements at time t are shown in Fig. 11.1.

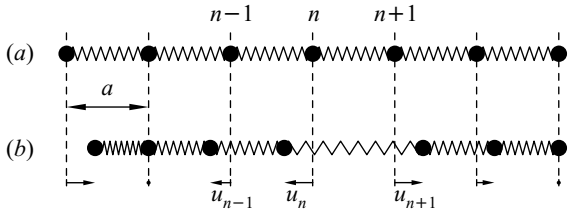


Fig. 11.1. Atomic positions in a one-dimensional monatomic chain (a) in equilibrium; (b) in vibration at an arbitrary time t. Springs represent the elastic forces between the atoms; u_n is the instantaneous displacement of the nth atom from its equilibrium position

Using the form (11.1.13) for the harmonic potential,

$$U_{\text{harm}} = \tfrac{1}{4} \sum_{nn'} \widetilde{\Phi}(n, n') \left[u_n - u_{n'}\right]^2 \tag{11.2.1}$$

in this simple case. It is plausible to assume that the strength of the interaction decreases rapidly with increasing separation of the atoms, and so it is sufficient to take into account only the effects of the nearest neighbors. Denoting the spring (force) constant $\widetilde{\Phi}(n, n \pm 1)$ by K,

$$U_{\text{harm}} = \tfrac{1}{2} K \sum_{n} \left[u_n - u_{n+1}\right]^2 . \tag{11.2.2}$$

The force on the atom is its derivative, i.e., the equation of motion for the nth atom is

$$M\ddot{u}_n = -\frac{\partial U_{\text{harm}}}{\partial u_n} = -K \left[2u_n - u_{n-1} - u_{n+1}\right] . \tag{11.2.3}$$

The solution of the system of coupled differential equations requires the specification of the boundary condition. A practical choice is the periodic or Born–von Kármán boundary condition, whereby the N atoms are assumed to be not along a free chain of length $L = Na$ but on a ring of the same circumference, and so the $N + 1$st atom is the same as the first. Then

$$u_{N+1} = u_1 . \tag{11.2.4}$$

As traveling waves are expected to propagate in the mass–spring system, the solution is most easily obtained using a Fourier expansion for the displacements. The discrete position variable is thus replaced by the discrete wave number q, and the continuous time variable by the continuous frequency variable ω:

$$u_n(t) = \frac{1}{\sqrt{N}} \sum_{q} \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{d}\omega \, u(q, \omega) \mathrm{e}^{\mathrm{i}(qna - \omega t)} . \tag{11.2.5}$$

Since displacements are real, the pairs (q, ω) and $(-q, -\omega)$ describe the same atomic vibration. Therefore we shall permit q but not ω to take negative values.

The periodic boundary condition allows only those values of q for which $\mathrm{e}^{\mathrm{i}qNa} = 1$, that is

$$q = j \frac{2\pi}{Na} , \qquad j = 0, \pm 1, \pm 2, \ldots . \tag{11.2.6}$$

When two integers j differ by an integral multiple of N, the difference of the corresponding wave numbers is an integral multiple of $2\pi/a$. As illustrated in Fig. 11.2, such waves describe the same atomic displacement because u_n is defined only in discrete lattice points.

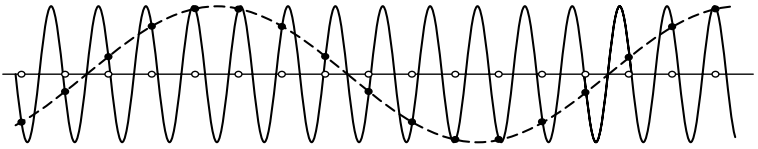


Fig. 11.2. Atomic displacements for the Fourier components with wave numbers q and $q + j\,2\pi/a$. For transparency, displacements are shown perpendicular to the chain. The graphs of the waves drawn with solid and dashed lines have physical meaning only in discrete lattice points

In the description of lattice vibrations wave numbers that differ by $2\pi/a$ are thus equivalent. Since $b = 2\pi/a$ is the primitive vector in the reciprocal lattice of a one-dimensional lattice, the previous assertion establishes the equivalence of wave vectors that differ by a reciprocal-lattice vector. This is in accordance with the general consequences of discrete translational symmetry discussed in Chapter 6. Once again, the N independent qs are customarily chosen in the Brillouin zone; for a one-dimensional chain this is the region

$$-\pi/a < q \leq \pi/a\,. \tag{11.2.7}$$

The restriction on wave numbers can be interpreted in another way by asserting that it is meaningless to speak about oscillations whose half wavelength is smaller than the lattice constant. Therefore we shall always represent the spectrum of lattice vibrations inside the Brillouin zone.

Substituting the Fourier transform into the equation of motion (11.2.3), the Fourier components of different wave numbers do not mix. In the harmonic approximation waves of different wavelengths propagate independently. The equation for the Fourier component $u(q,\omega)$ is

$$\begin{aligned} -M\omega^2 u(q,\omega) &= -K\left[2 - \mathrm{e}^{-\mathrm{i}qa} - \mathrm{e}^{\mathrm{i}qa}\right] u(q,\omega) \\ &= -2K\left[1 - \cos qa\right] u(q,\omega)\,. \end{aligned} \tag{11.2.8}$$

The angular frequency of the oscillation of wave number q is then

$$\omega(q) = \sqrt{\frac{2K(1-\cos qa)}{M}} = 2\left(\frac{K}{M}\right)^{1/2} \left|\sin\left(\tfrac{1}{2}qa\right)\right|\,. \tag{11.2.9}$$

The relationship between the vibrational frequency[1] and the wave number – the dispersion relation – is shown in Fig. 11.3.

In the long-wavelength limit ($1/q \gg a$) the curve is linear:

$$\omega(q) \approx \left(\frac{K}{M}\right)^{1/2} a\,|q|\,. \tag{11.2.10}$$

[1] The notation ω is always used for angular frequencies, however we shall follow common practice and use the term *frequency* instead of *angular frequency*.

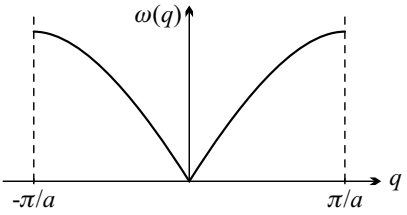


Fig. 11.3. Dispersion curve for the longitudinal vibrations of a monatomic linear chain

In this limit neighboring atoms oscillate almost perfectly in phase, and the motion corresponds to the oscillation of an elastic continuum, i.e., the propagation of sound waves. For this reason such vibrations are called *acoustic vibrations*.

When lattice vibrations are combined into a wave packet, the group velocity is

$$c = \left| \frac{\partial \omega(q)}{\partial q} \right| = \left(\frac{K}{M} \right)^{1/2} a \, \cos \frac{qa}{2} \, . \tag{11.2.11}$$

Starting from the midpoint of the Brillouin zone the group velocity is monotonically decreasing. As shown in the figure, the dispersion curve has extrema at the boundaries of the Brillouin zone; it becomes flat there, so its derivative, the group velocity vanishes. This is the consequence of the reflection symmetry in monatomic linear chains.

It is straightforward to demonstrate that some characteristic features of the dispersion curve are preserved when not only first neighbors interact. Denoting the force constant of the interaction between pth neighbors by K_p, the potential is

$$U_{\text{harm}} = \tfrac{1}{2} \sum_{n,p} K_p \left[u_n - u_{n+p} \right]^2 \tag{11.2.12}$$

in the harmonic approximation, and so the equation of motion is

$$M \ddot{u}_n = \sum_p K_p \left[2u_n - u_{n+p} - u_{n-p} \right] . \tag{11.2.13}$$

Eigenfrequencies are then given by

$$\omega^2(q) = \frac{2}{M} \sum_p K_p \left[1 - \cos(pqa) \right] \, . \tag{11.2.14}$$

It is readily seen that ω is always an even function of q,

$$\omega(q) = \omega(-q) \, . \tag{11.2.15}$$

This formula reflects the symmetry that waves can equally propagate to the left and to the right along the chain. This result is generally valid: as it was

mentioned in Chapter 6 it is the consequence of invariance under time reversal. When interactions are not limited to nearest neighbors, the dispersion curve still starts linearly, and the group velocity still vanishes at the boundary of the Brillouin zone.

Up to now it has been assumed that atomic displacements are along the chain. Such vibrations are called *longitudinal*. For each q there exists one such solution. When atomic displacements are perpendicular to the chain, we speak of *transverse vibrations*. As there are two perpendicular directions, two transverse vibrations are associated with each value of q. In a lattice made up of N atoms $3N$ vibrational states are therefore possible. However, in a linear chain this is true only in principle. Displacements perpendicular to the chain modify the energy to a lesser extent than parallel ones. As mentioned in connection with (11.1.19), transverse displacements are not opposed by restoring forces in the harmonic approximation, consequently transverse vibrations cannot propagate in the linear chain. This is no longer the case in two- and three-dimensional crystals. Here transverse vibrations are of finite frequency, too, provided certain atomic bonds are not in the plane spanned by the propagation direction and the direction of vibration.

11.2.2 Vibrations of a Diatomic Chain

The vibrational spectrum is more complicated when the primitive cell of the one-dimensional chain contains two atoms. There are two limits of particular interest. In the first an atom of mass M_1 is located at the lattice point and another of mass M_2 at the midpoint of the cell.

We shall denote the displacement from equilibrium of the atom of mass M_1 (M_2) in the nth primitive cell by u_n (v_n). Equilibrium positions and instantaneous positions at a given time are shown in Fig. 11.4.

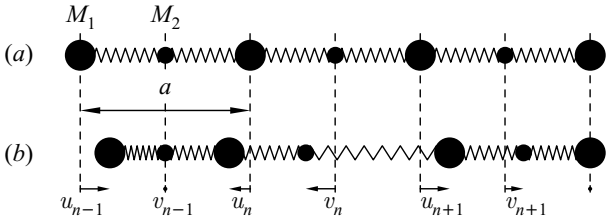


Fig. 11.4. Atomic positions in the primitive cell of a linear chain made up of two kinds of atom. (a) Equilibrium positions; (b) instantaneous displacements

Assuming that each atom feels the potential arising from its two nearest neighbors, the system can be characterized by a single force constant K. The potential energy is

$$U_{\text{harm}} = \tfrac{1}{2}K\sum_n [u_n - v_n]^2 + \tfrac{1}{2}K\sum_n [v_n - u_{n+1}]^2. \tag{11.2.16}$$

Then the equations of motion for the two kinds of atom are

$$\begin{aligned} M_1\ddot{u}_n &= -K\left[2u_n - v_n - v_{n-1}\right], \\ M_2\ddot{v}_n &= -K\left[2v_n - u_{n+1} - u_n\right]. \end{aligned} \tag{11.2.17}$$

From the results obtained for the vibrations in a monatomic linear chain the solutions are expected to be linear combinations of independently propagating waves. Choosing a component of wave number q and frequency ω, the amplitudes are assumed to be different on the two sublattices:

$$u_n(t) = u(q)\mathrm{e}^{\mathrm{i}(qna-\omega t)}\,, \qquad v_n(t) = v(q)\mathrm{e}^{\mathrm{i}(qna-\omega t)}\,. \tag{11.2.18}$$

Using periodic boundary conditions means that the requirement

$$u_{N+1} = u_1\,, \qquad v_{N+1} = v_1 \tag{11.2.19}$$

applies, and the chain is closed into a ring made up of N primitive cells. The allowed values for q are the same as for the monatomic chain.

Substituting the traveling wave form into the equations of motion, elimination of the common phase factor yields

$$\begin{aligned} -\omega^2 M_1 u(q) &= -2Ku(q) + Kv(q)(1+\mathrm{e}^{-\mathrm{i}qa})\,, \\ -\omega^2 M_2 v(q) &= -2Kv(q) + Ku(q)(\mathrm{e}^{\mathrm{i}qa}+1)\,. \end{aligned} \tag{11.2.20}$$

For this equation to have a nontrivial solution the determinant of the coefficient matrix has to vanish:

$$\begin{vmatrix} -\omega^2 M_1 + 2K & -2K\mathrm{e}^{-\mathrm{i}qa/2}\cos\left(\frac{1}{2}qa\right) \\ -2K\mathrm{e}^{\mathrm{i}qa/2}\cos\left(\frac{1}{2}qa\right) & -\omega^2 M_2 + 2K \end{vmatrix} = 0\,. \tag{11.2.21}$$

The vibration frequencies are then

$$\omega_\pm^2(q) = \frac{K}{M_1 M_2}\left\{(M_1+M_2) \pm \sqrt{(M_1+M_2)^2 - 4M_1M_2\sin^2\left(\tfrac{1}{2}qa\right)}\right\}. \tag{11.2.22}$$

Using the notation

$$\omega_0^2 = 2K\left(\frac{1}{M_1}+\frac{1}{M_2}\right), \qquad \gamma^2 = 4\frac{M_1M_2}{(M_1+M_2)^2}\,, \tag{11.2.23}$$

they can be written as

$$\omega_\pm^2(q) = \tfrac{1}{2}\omega_0^2\left\{1 \pm \sqrt{1-\gamma^2\sin^2\left(\tfrac{1}{2}qa\right)}\right\}. \tag{11.2.24}$$

In contrast to the single longitudinal mode in the monatomic chain there are now two such modes with different frequencies. The corresponding dispersion relations are shown in Fig. 11.5.

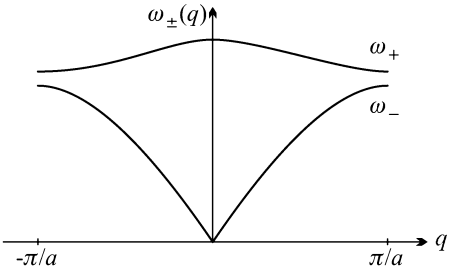


Fig. 11.5. Dispersion relations in the acoustic (ω_-) and optical (ω_+) branches for the vibrations in a diatomic chain

The dispersion relation of one type of vibration ($\omega_-(q)$) is very similar to that of the acoustic vibrations in a monatomic chain. It vanishes for $q = 0$, and grows linearly for small q:

$$\omega_-(q) \approx \tfrac{1}{4}\omega_0\gamma|q|\,a\,. \tag{11.2.25}$$

The other branch of the dispersion relation starts at a finite frequency:

$$\omega_+(q=0) = \omega_0\,. \tag{11.2.26}$$

The ratio of the amplitudes $u(q)$ and $v(q)$ is determined from (11.2.20). For $q = 0$, i.e., at the center of the Brillouin zone

$$\frac{v(0)}{u(0)} = \begin{cases} +1 & \text{in the branch } \omega_-, \\ -(M_1/M_2) & \text{in the branch } \omega_+. \end{cases} \tag{11.2.27}$$

In the branch with frequency ω_- the two atoms of the primitive cell oscillate with almost equal amplitudes and are in phase in the long-wavelength limit – just like in a sound wave. For this reason this lower branch of vibrations is called acoustic branch in this case, too. In the branch with frequency ω_+ the two atoms oscillate in opposite directions around their center of mass in the long-wavelength limit. Oscillation amplitudes on the two sublattices are inversely proportional to the masses. When the diatomic chain is an ionic crystal made up of oppositely charged ions, such vibrations may be excited by high-frequency electromagnetic fields (light); they are therefore called *optical vibrations*.

In the acoustic branch, at the boundary of the Brillouin zone

$$\begin{aligned} \omega_-(q) &= \sqrt{K}\left[\left(\frac{1}{M_1}+\frac{1}{M_2}\right) - \left|\frac{1}{M_1}-\frac{1}{M_2}\right|\right]^{1/2} \\ &= \begin{cases} (2K/M_1)^{1/2}\,, & \text{if} \quad M_1 > M_2\,, \\ (2K/M_2)^{1/2}\,, & \text{if} \quad M_2 > M_1\,. \end{cases} \end{aligned} \tag{11.2.28}$$

The situation is just the opposite in the optical branch. Here the frequency is $(2K/M_2)^{1/2}$ for $M_1 > M_2$ and $(2K/M_1)^{1/2}$ for $M_2 > M_1$ at the boundary of the Brillouin zone. Consequently there is always a finite gap between the two branches: the acoustic branch is always below the optical. Examined as functions of q, both branches of excitation are flat at the boundary of the Brillouin zone, so the group velocity vanishes there.

Note that at the boundary of the Brillouin zone the frequency of the acoustic (optical) branch depends only on the mass of the heavier (lighter) atom. This can be easily understood by taking the ratio of the two vibration amplitudes. At the boundary of the Brillouin zone it is either zero or infinity – implying that only one type of atom participates in either vibration. Figure 11.6 shows the atomic displacements in the acoustic and optical branches for wave numbers at the center and boundary of the Brillouin zone.

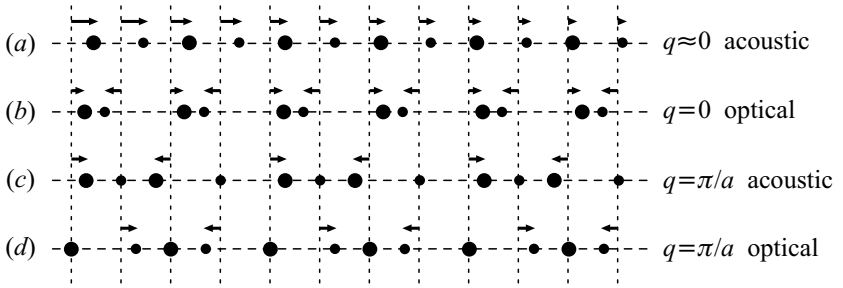


Fig. 11.6. Atomic displacements in the acoustic and optical modes of a diatomic linear chain in the long-wavelength limit ($q \approx 0$), and for a wave number at the zone boundary

Let us examine what happens when the two masses are changed continuously and become equal. Denoting the common mass by M, (11.2.22) implies

$$\omega_{\pm}^2(q) = \frac{2K}{M}\left[1 \pm \cos\left(\tfrac{1}{2}qa\right)\right], \tag{11.2.29}$$

that is,

$$\omega_{\pm}(q) = \begin{cases} 2\,(K/M)^{1/2}\,|\cos\left(\tfrac{1}{4}qa\right)|\,, \\ 2\,(K/M)^{1/2}\,|\sin\left(\tfrac{1}{4}qa\right)|\,. \end{cases} \tag{11.2.30}$$

The dispersion relation valid for this special case is shown in Fig. 11.7.

Both the figure and the analytical expressions show that the dispersion curves are now not perpendicular to the zone boundary at π/a. This is not surprising as we are now dealing with a chain that contains $2N$ equivalent atoms spaced uniformly over a length Na, i.e., separated by regular distances $a/2$. Therefore the chain is in fact monatomic, and the dimension of its true primitive cell is $a/2$. The boundary of the Brillouin zone is then at $2\pi/a$

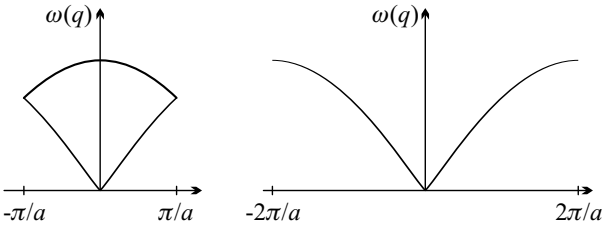


Fig. 11.7. Dispersion relation for the vibrations of a diatomic linear chain in the $M_1 = M_2$ limit, shown in the Brillouin zones of chains with lattice constants a and $a/2$

instead of π/a. In this "large" Brillouin zone there are $2N$ allowed values for q, each associated with one eigenfrequency – while if the Brillouin zone that corresponds to a periodicity a is used, the number of allowed qs is N, and each of them is associated with two frequencies. In the $M_1 = M_2$ limit the optical vibrations obtained in the diatomic chain correspond to those acoustic vibrations of a monatomic chain for which the wave number is either $\pi/a < q < -\pi/a$ or $\pi/a < q < 2\pi/a$. Shifting the optical branch by $\pm 2\pi/a$ into these intervals, the dispersion curve of the monatomic chain is recovered.

11.2.3 Vibrations of a Dimerized Chain

In the other limit the two atoms in the primitive cell are of equal mass but are not uniformly spaced along the chain: the separation between nearest neighbors alternates regularly between a smaller and a larger value. Such a configuration – illustrated in Fig. 11.8 – is called a *dimerized chain*.

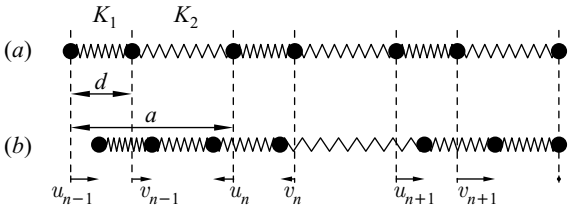


Fig. 11.8. Equilibrium positions and instantaneous displacements of the atoms in a dimerized chain

Let the equilibrium atomic positions in the nth cell be denoted by na and $na + d$, and the displacements from them by u_n and v_n. Depending on whether neighboring atoms are separated by d or $a-d$, the pair potential takes different values. For this reason even when only nearest-neighbor interactions are taken into account, two force constants have to be introduced. In the

harmonic approximation the expression for the potential reads

$$U_{\text{harm}} = \frac{K_1}{2}\sum_n \left[u_n - v_n\right]^2 + \frac{K_2}{2}\sum_n \left[v_n - u_{n+1}\right]^2 . \tag{11.2.31}$$

When $d \le a/2$, it is plausible to assume that $K_1 \ge K_2$.

Determining the force on the atoms from the energy expression, the equations of motion are

$$\begin{aligned} M\ddot{u}_n &= -\frac{\partial U_{\text{harm}}}{\partial u_n} = -K_1\left[u_n - v_n\right] - K_2\left[u_n - v_{n-1}\right], \\ M\ddot{v}_n &= -\frac{\partial U_{\text{harm}}}{\partial v_n} = -K_1\left[v_n - u_n\right] - K_2\left[v_n - u_{n+1}\right]. \end{aligned} \tag{11.2.32}$$

Traveling-wave solutions are sought in the form

$$u_n(t) = u(q)\mathrm{e}^{\mathrm{i}(qna-\omega t)}, \qquad v_n(t) = v(q)\mathrm{e}^{\mathrm{i}(qna-\omega t)}; \tag{11.2.33}$$

substitution into the equations of motion then gives

$$\begin{aligned} -\omega^2 M u(q) &= -(K_1+K_2)u(q) + \left(K_1 + K_2\mathrm{e}^{-\mathrm{i}qa}\right)v(q), \\ -\omega^2 M v(q) &= -(K_1+K_2)v(q) + \left(K_1 + K_2\mathrm{e}^{\mathrm{i}qa}\right)u(q). \end{aligned} \tag{11.2.34}$$

Nontrivial solutions exist if the determinant of the coefficients of $u(q)$ and $v(q)$ vanishes, i.e.,

$$\begin{aligned} \left[-\omega^2 M + (K_1+K_2)\right]^2 &= \left(K_1 + K_2\mathrm{e}^{-\mathrm{i}qa}\right)\left(K_1 + K_2\mathrm{e}^{\mathrm{i}qa}\right) \\ &= K_1^2 + K_2^2 + 2K_1K_2\cos qa . \end{aligned} \tag{11.2.35}$$

From this equation the allowed frequencies are

$$\omega_\pm^2(q) = \frac{1}{M}\left\{(K_1+K_2) \pm \sqrt{K_1^2 + K_2^2 + 2K_1K_2\cos qa}\right\}, \tag{11.2.36}$$

and the amplitude ratio is

$$\frac{v(q)}{u(q)} = \mp\sqrt{\frac{K_1 + K_2\mathrm{e}^{\mathrm{i}qa}}{K_1 + K_2\mathrm{e}^{-\mathrm{i}qa}}}. \tag{11.2.37}$$

Vibrational frequencies can again be written in the form of (11.2.24) with

$$\omega_0^2 = (K_1+K_2)\frac{2}{M}, \qquad \gamma^2 = 4\frac{K_1K_2}{(K_1+K_2)^2}. \tag{11.2.38}$$

Similarly to diatomic chains, the spectrum has two branches. The lower branch starts linearly,

$$\omega_-(q) \approx \tfrac{1}{4}\omega_0\gamma|q|\,a, \tag{11.2.39}$$

and since for q small $u(q) \approx v(q)$, atoms oscillate almost in phase in the long-wavelength limit. This is the acoustic branch.

In contrast, the upper branch starts at a nonzero frequency ω_0 at the center of the Brillouin zone,

$$\omega_+(q) = \omega_0 - \mathcal{O}(qa)^2 \,. \tag{11.2.40}$$

Since for q small $v(q) \approx -u(q)$ now, the two atoms oscillate in opposite phases. This branch is called the optical branch here, too.

At the boundary of the Brillouin zone, for $q = \pi/a$

$$\omega_\pm^2(\pi/a) = \frac{1}{M}\left[(K_1 + K_2) \pm |K_1 - K_2|\right] . \tag{11.2.41}$$

For $K_1 > K_2$ this gives

$$\omega_+(\pi/a) = (2K_1/M)^{1/2} \,, \qquad \omega_-(\pi/a) = (2K_2/M)^{1/2} \,. \tag{11.2.42}$$

The ratio of the amplitudes of the displacements is found to be ∓ 1. This means that atomic displacements are such that atoms separated by d in one branch and by $a - d$ in the other branch oscillate in phase – and so only one of two springs is stretched in either branch.

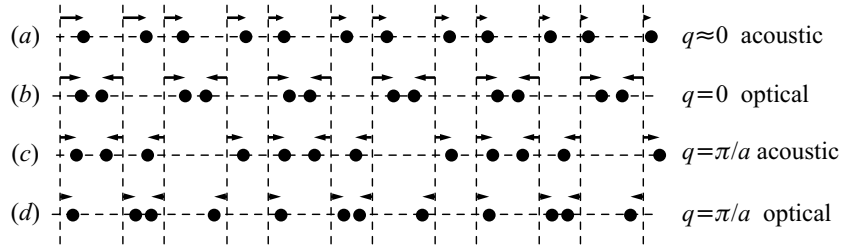


Fig. 11.9. Atomic displacements in the acoustic and optical modes of a dimerized linear chain in the long-wavelength limit for $q = \pi/a$

In the limit $K_1 \gg K_2$,

$$\begin{aligned} \omega_-(q) &= (2K_2/M)^{1/2} \left|\sin(qa/2)\right| \left[1 + \mathcal{O}(K_2/K_1)\right] , \qquad && v(q) \approx u(q) \,, \\ \omega_+(q) &= (2K_1/M)^{1/2} \left[1 + \mathcal{O}(K_2/K_1)\right] , && v(q) \approx -u(q) \,. \end{aligned} \tag{11.2.43}$$

Comparison with the vibrational frequencies (11.2.9) shows that the acoustic branch is just like for a monatomic linear chain with atoms of mass $2M$ and springs of force constant K_2. This can be interpreted as follows: each strong spring of force constant K_1 binds two atoms into an almost perfectly rigid "molecule" of mass $2M$, and the chain made up of such "molecules"

produces acoustic vibrations. On the other hand, each vibration in the optical branch has practically the same frequency, regardless of the wavelength. This vibrational frequency is the frequency of internal oscillations of a "diatomic molecule", made up of two atoms of mass M and held together by a spring of force constant K_1.

In the more general case, when the dimerized chain is built up of two atoms of unequal mass, the equations of motion for the two atoms are

$$\begin{aligned} M_1\ddot{u}_n &= -K_1\left[u_n - v_n\right] - K_2\left[u_n - v_{n-1}\right], \\ M_2\ddot{v}_n &= -K_1\left[v_n - u_n\right] - K_2\left[v_n - u_{n+1}\right]. \end{aligned} \tag{11.2.44}$$

Seeking traveling-wave solutions, the amplitudes need to satisfy the homogeneous linear system of equations

$$\begin{aligned} &(-\omega^2 M_1 + K_1 + K_2)u(q) - (K_1 + K_2\mathrm{e}^{-\mathrm{i}qa})v(q) = 0\,, \\ &-(K_1 + K_2\mathrm{e}^{\mathrm{i}qa})u(q) + (-\omega^2 M_2 + K_1 + K_2)v(q) = 0\,. \end{aligned} \tag{11.2.45}$$

Nontrivial solutions exist when the determinant of coefficients is zero; the frequencies of the vibrations are then written as

$$\omega_{\pm}^2(q) = \tfrac{1}{2}\omega_0^2\left\{1 \pm \sqrt{1 - \gamma^2\sin^2\left(\tfrac{1}{2}qa\right)}\right\}, \tag{11.2.46}$$

where

$$\omega_0^2 = (K_1 + K_2)\left(\frac{1}{M_1} + \frac{1}{M_2}\right) \tag{11.2.47}$$

and

$$\gamma^2 = 16\frac{K_1K_2}{(K_1 + K_2)^2}\frac{M_1M_2}{(M_1 + M_2)^2}\,. \tag{11.2.48}$$

In the $K_1 = K_2$ ($M_1 = M_2$) limit the results derived for diatomic (dimerized) chains are recovered.

In the general case an acoustic branch starting off at zero frequency and an optical branch starting off at some finite frequency are found. The vibrational frequencies at the center and boundary of the Brillouin zone are

$$\omega_-(q) = \begin{cases} \frac{1}{4}\omega_0\gamma|q|a & \text{at } q \to 0, \\ \dfrac{1}{\sqrt{2}}\omega_0\sqrt{1 - \sqrt{1 - \gamma^2}} & \text{at } q = \pm\pi/a, \end{cases} \tag{11.2.49}$$

in the acoustic branch and

$$\omega_-(q) = \begin{cases} \omega_0 & \text{at } q = 0, \\ \dfrac{1}{\sqrt{2}}\omega_0\sqrt{1 + \sqrt{1 - \gamma^2}} & \text{at } q = \pm\pi/a. \end{cases} \tag{11.2.50}$$

in the optical branch.

11.2.4 Vibrations of a Simple Cubic Lattice

Having examined some one-dimensional examples, let us now turn to the vibrational spectrum of simple cubic crystals with a monatomic basis, employing the approximation that the only first- and second-neighbor interactions contribute to the potential.

Lattice points will be specified by their coordinate indices, and for each point (i,j,k) the six first neighbors $(i\pm1,j,k)$, $(i,j\pm1,k)$, $(i,j,k\pm1)$ and the twelve second neighbors $(i\pm1,j\pm1,k)$, $(i\pm1,j,k\pm1)$, $(i,j\pm1,k\pm1)$ will be taken into account. Calculations are highly simplified by the remark made in connection with (11.1.19): in the harmonic approximation restoring forces arise only for displacements along the line joining the atoms. For displacements in the perpendicular direction the change in the length of the spring is of second order, and can therefore be neglected. This way only two force constants remain: one characterizes the change in energy due to the relative displacement of nearest-neighbor atoms along the line joining them, and the second is related to the change in energy due to the relative displacement of second-neighbor atoms along the face diagonal:

$$
\begin{aligned}
U_{\text{harm}} = {} & \tfrac{1}{2}K_1 \sum_{ijk} \Big\{ [u_x(i,j,k) - u_x(i+1,j,k)]^2 \\
& + [u_y(i,j,k) - u_y(i,j+1,k)]^2 + [u_z(i,j,k) - u_z(i,j,k+1)]^2 \Big\} \\
& + \tfrac{1}{4}K_2 \sum_{ijk} \Big\{ [u_x(i,j,k) - u_x(i+1,j+1,k) \\
& \qquad + u_y(i,j,k) - u_y(i+1,j+1,k)]^2 \\
& + [u_x(i,j,k) - u_x(i+1,j-1,k) - u_y(i,j,k) + u_y(i+1,j-1,k)]^2 \\
& + [u_x(i,j,k) - u_x(i+1,j,k+1) + u_z(i,j,k) - u_z(i+1,j,k+1)]^2 \\
& + [u_x(i,j,k) - u_x(i+1,j,k-1) - u_z(i,j,k) + u_z(i+1,j,k-1)]^2 \\
& + [u_y(i,j,k) - u_y(i,j+1,k+1) + u_z(i,j,k) - u_z(i,j+1,k+1)]^2 \\
& + [u_y(i,j,k) - u_y(i,j+1,k-1) - u_z(i,j,k) + u_z(i,j+1,k-1)]^2 \Big\}.
\end{aligned}
\tag{11.2.51}
$$

The force on the atom sitting at the lattice point labeled (i,j,k) is the derivative of the potential:

$$
\begin{aligned}
F_x(i,j,k) &= -\frac{\partial U_{\text{harm}}}{\partial u_x(i,j,k)} \\
&= K_1\left[2u_x(i,j,k) - u_x(i+1,j,k) - u_x(i-1,j,k)\right] \\
&\quad +\tfrac{1}{2}K_2\big[2u_x(i,j,k) - u_x(i+1,j+1,k) - u_x(i-1,j-1,k) \\
&\qquad + 2u_x(i,j,k) - u_x(i+1,j-1,k) - u_x(i-1,j+1,k) \\
&\qquad + 2u_x(i,j,k) - u_x(i+1,j,k+1) - u_x(i-1,j,k-1) \\
&\qquad + 2u_x(i,j,k) - u_x(i+1,j,k-1) - u_x(i-1,j,k+1) \\
&\qquad - u_y(i+1,j+1,k) - u_y(i-1,j-1,k) \\
&\qquad + u_y(i+1,j-1,k) + u_y(i-1,j+1,k) \\
&\qquad - u_z(i+1,j,k+1) - u_z(i-1,j,k-1) \\
&\qquad + u_z(i+1,j,k-1) + u_z(i-1,j,k+1)\big].
\end{aligned}
\tag{11.2.52}
$$

The y- and z-components of the force are given by similar expressions. Substituting them into the equation of motion, and seeking traveling-wave solutions,

$$
u_\alpha(\boldsymbol{R},t) = u_\alpha(\boldsymbol{q})\exp\{\mathrm{i}(q_xR_x + q_yR_y + q_zR_z - \omega t)\}\,. \tag{11.2.53}
$$

Using the Born–von Kármán boundary conditions, the allowed values of q are once again expressed in terms of the primitive vectors of the reciprocal lattice as

$$
\boldsymbol{q} = \frac{n_1}{N_1}\boldsymbol{b}_1 + \frac{n_2}{N_2}\boldsymbol{b}_2 + \frac{n_3}{N_3}\boldsymbol{b}_3\,, \tag{11.2.54}
$$

where n_1, n_2, and n_3 are integers. Two vectors $\boldsymbol{q}$ that differ by a reciprocal-lattice vector describe the same atomic displacement, consequently there are only $N = N_1\,N_2\,N_3$ physically different vectors $\boldsymbol{q}$. Here, too, it is useful to choose them inside the Brillouin zone.

The coefficients $u_\alpha(\boldsymbol{q})$ are determined by the homogeneous linear system of equations

$$
\sum_\beta D_{\alpha\beta}(\boldsymbol{q})u_\beta(\boldsymbol{q}) = \omega^2 u_\alpha(\boldsymbol{q})\,, \tag{11.2.55}
$$

where

$$
\begin{aligned}
D_{xx}(\boldsymbol{q}) &= \frac{2K_1}{M}(1-\cos q_xa) + \frac{2K_2}{M}(2-\cos q_xa\cos q_ya - \cos q_xa\cos q_za)\,, \\
D_{yy}(\boldsymbol{q}) &= \frac{2K_1}{M}(1-\cos q_ya) + \frac{2K_2}{M}(2-\cos q_ya\cos q_xa - \cos q_ya\cos q_za)\,, \\
D_{zz}(\boldsymbol{q}) &= \frac{2K_1}{M}(1-\cos q_za) + \frac{2K_2}{M}(2-\cos q_za\cos q_xa - \cos q_za\cos q_ya)\,, \\
D_{xy}(\boldsymbol{q}) &= D_{yx}(\boldsymbol{q}) = \frac{2K_2}{M}\sin q_xa\sin q_ya\,, \\
D_{xz}(\boldsymbol{q}) &= D_{zx}(\boldsymbol{q}) = \frac{2K_2}{M}\sin q_xa\sin q_za\,, \\
D_{yz}(\boldsymbol{q}) &= D_{zy}(\boldsymbol{q}) = \frac{2K_2}{M}\sin q_ya\sin q_za\,.
\end{aligned}
\tag{11.2.56}
$$

Note that (11.2.55) can also be obtained by Fourier transforming (11.1.29), provided $D_{\alpha\beta}(\boldsymbol{q})$ is defined as

$$D_{\alpha\beta}(\boldsymbol{q}) = \frac{1}{M} \sum_{m \neq n} \left(1 - \mathrm{e}^{-\mathrm{i}\boldsymbol{q}\cdot(\boldsymbol{R}_m - \boldsymbol{R}_n)}\right) \widetilde{\Phi}_{\alpha\beta}(m-n) \tag{11.2.57}$$

when the basis consists of a single atom. The comparison of (11.1.13) and (11.2.51) immediately yields $\widetilde{\Phi}_{\alpha\beta}(m-n)$, and then the force constants can be obtained directly from the previous equation.

To determine the eigenfrequencies, the eigenvalues of the 3×3 matrix $\mathsf{D}(\boldsymbol{q})$ have to be calculated. This requires the solution of a cubic equation for ω^2, which gives three different frequencies. In special cases, in high-symmetry points of the Brillouin zone the eigenvalues may nevertheless become degenerate.

The Brillouin zone of the simple cubic crystal is shown in Fig. 7.2. For vibrations propagating along the direction [100] and characterized by the wave vector $\Delta = (q, 0, 0)$ (that lies along the line between Γ and X) the matrix $\mathsf{D}(\boldsymbol{q})$ is diagonal:

$$\mathsf{D}(q,0,0) = \begin{pmatrix} A & 0 & 0 \\ 0 & B & 0 \\ 0 & 0 & B \end{pmatrix}, \tag{11.2.58}$$

where

$$\begin{aligned} A &= \frac{2K_1 + 4K_2}{M}(1 - \cos qa) = \frac{4K_1 + 8K_2}{M} \sin^2\left(\tfrac{1}{2}qa\right), \\ B &= \frac{2K_2}{M}(1 - \cos qa) = \frac{4K_2}{M} \sin^2\left(\tfrac{1}{2}qa\right). \end{aligned} \tag{11.2.59}$$

It is immediately recognized that the crystal has a nondegenerate and a doubly degenerate eigenfrequency:

$$\omega_1 = \sqrt{\frac{4K_1 + 8K_2}{M}} \left|\sin\left(\tfrac{1}{2}qa\right)\right|, \qquad \omega_{2,3} = \sqrt{\frac{4K_2}{M}} \left|\sin\left(\tfrac{1}{2}qa\right)\right|. \tag{11.2.60}$$

In the harmonic approximation the amplitude of the vibrations can be arbitrarily large, only their directions are determined by the eigenvectors of the matrix $\mathsf{D}(\boldsymbol{q})$. The eigenvectors for the above eigenfrequencies are

$$\boldsymbol{e}^{(1)} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad \boldsymbol{e}^{(2)} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \qquad \boldsymbol{e}^{(3)} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{11.2.61}$$

The first solution describes atoms moving in the propagation direction of the wave; the vibration is longitudinal in this case. For the two other solutions atomic displacements are perpendicular to the propagation direction; these vibrations are therefore transverse. Note that the condition for the existence

of transverse waves is that K_2 be finite – that is, in a simple cubic crystal force constants should be nonzero not only along the direction of propagation or perpendicular to it but also in the diagonal direction.

The propagation velocities for the two waves are

$$c_{\mathrm{L}} = a\left(\frac{K_1 + 2K_2}{M}\right)^{1/2}, \qquad c_{\mathrm{T}} = a\left(\frac{K_2}{M}\right)^{1/2}. \tag{11.2.62}$$

The matrix is diagonal for the points $Z = (\pi/a, q, 0)$ along the line between X and M in Fig. 7.2, too:

$$\boldsymbol{D}(\pi/a, q, 0) = \begin{pmatrix} A & 0 & 0 \\ 0 & B & 0 \\ 0 & 0 & C \end{pmatrix}, \tag{11.2.63}$$

where

$$\begin{aligned} A &= \frac{4K_1}{M} + \frac{2K_2}{M}(3 + \cos qa) = \frac{4K_1 + 8K_2}{M} - \frac{4K_2}{M}\sin^2\left(\tfrac{1}{2}qa\right), \\ B &= \frac{2K_1}{M}(1 - \cos qa) + \frac{4K_2}{M} = \frac{4K_2}{M} + \frac{4K_1}{M}\sin^2\left(\tfrac{1}{2}qa\right), \\ C &= \frac{2K_2}{M}(3 - \cos qa) = \frac{4K_2}{M} + \frac{4K_2}{M}\sin^2\left(\tfrac{1}{2}qa\right). \end{aligned} \tag{11.2.64}$$

The eigenvectors are the same as those given in (11.2.61), and the eigenfrequencies are given by the square root of A, B, and C. The vibration characterized by the eigenvector $\boldsymbol{e}^{(3)}$ is transverse, while in the two other cases the eigenvector $\boldsymbol{e}$ is neither longitudinal nor transverse. This clearly indicates that in the general case one cannot speak of purely longitudinal and purely transverse oscillations.

The matrix is no longer diagonal for the points $\Sigma = (q, q, 0)$ along the line between M and X:

$$\boldsymbol{D}(q, q, 0) = \begin{pmatrix} A & B & 0 \\ B & A & 0 \\ 0 & 0 & C \end{pmatrix}, \tag{11.2.65}$$

where

$$\begin{aligned} A &= \frac{2K_1 + 2K_2}{M}(1 - \cos qa) + \frac{2K_2}{M}(1 - \cos^2 qa) \\ &= \frac{4K_1 + 4K_2}{M}\sin^2\left(\tfrac{1}{2}qa\right) + \frac{2K_2}{M}\sin^2 qa, \\ B &= \frac{2K_2}{M}\sin^2 qa, \\ C &= \frac{4K_2}{M}(1 - \cos qa) = \frac{8K_2}{M}\sin^2\left(\tfrac{1}{2}qa\right). \end{aligned} \tag{11.2.66}$$

It is an elementary exercise to find the eigenvectors of the matrix:

$$\boldsymbol{e}^{(1)} = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\\0\end{pmatrix}, \qquad \boldsymbol{e}^{(2)} = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\\0\end{pmatrix}, \qquad \boldsymbol{e}^{(3)} = \begin{pmatrix}0\\0\\1\end{pmatrix}. \tag{11.2.67}$$

The corresponding eigenfrequencies are

$$\begin{aligned}
\omega_1 &= \sqrt{A+B} = \sqrt{\frac{4K_1+4K_2}{M}\sin^2\left(\tfrac{1}{2}qa\right) + \frac{4K_2}{M}\sin^2 qa}\,,\\
\omega_2 &= \sqrt{A-B} = \sqrt{\frac{4K_1+4K_2}{M}}\left|\sin\left(\tfrac{1}{2}qa\right)\right|,\\
\omega_3 &= \sqrt{C} = \sqrt{\frac{8K_2}{M}}\left|\sin\left(\tfrac{1}{2}qa\right)\right|.
\end{aligned} \tag{11.2.68}$$

The highest-frequency vibration (given by the first solution) is longitudinal; the two others are transverse.

Finally, for vibrations propagating along the direction [111] and characterized by the wave vectors $\Lambda = (q,q,q)$ the matrix $\mathsf{D}(\boldsymbol{q})$ is of the form

$$\mathsf{D}(q,q,q) = \begin{pmatrix}A & B & B\\ B & A & B\\ B & B & A\end{pmatrix}, \tag{11.2.69}$$

where

$$\begin{aligned}
A &= \frac{2K_1}{M}(1-\cos qa) + \frac{4K_2}{M}(1-\cos^2 qa) = \frac{4K_1}{M}\sin^2\left(\tfrac{1}{2}qa\right) + \frac{4K_2}{M}\sin^2 qa\,,\\
B &= \frac{2K_2}{M}(1-\cos^2 qa) = \frac{2K_2}{M}\sin^2 qa\,.
\end{aligned} \tag{11.2.70}$$

This matrix is also straightforward to diagonalize. The eigenvectors are

$$\boldsymbol{e}^{(1)} = \frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\1\end{pmatrix}, \quad \boldsymbol{e}^{(2)} = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\\0\end{pmatrix}, \quad \boldsymbol{e}^{(3)} = \frac{1}{\sqrt{6}}\begin{pmatrix}1\\1\\-2\end{pmatrix}, \tag{11.2.71}$$

and the corresponding eigenfrequencies are

$$\begin{aligned}
\omega_1 &= \sqrt{A+2B} = \sqrt{\frac{4K_1}{M}\sin^2\left(\tfrac{1}{2}qa\right) + \frac{8K_2}{M}\sin^2 qa}\,,\\
\omega_{2,3} &= \sqrt{A-B} = \sqrt{\frac{4K_1}{M}\sin^2\left(\tfrac{1}{2}qa\right) + \frac{2K_2}{M}\sin^2 qa}\,.
\end{aligned} \tag{11.2.72}$$

Once again, a longitudinal and two degenerate transverse vibrations are found. Figure 11.10 shows the dispersion relations for the lattice vibrations along the four characteristic directions of the Brillouin zone discussed above.

In this crystal structure three acoustic branches are observed; they start from zero frequency at $\boldsymbol{q} = 0$. For wave vectors $\boldsymbol{q}$ that are along directions of

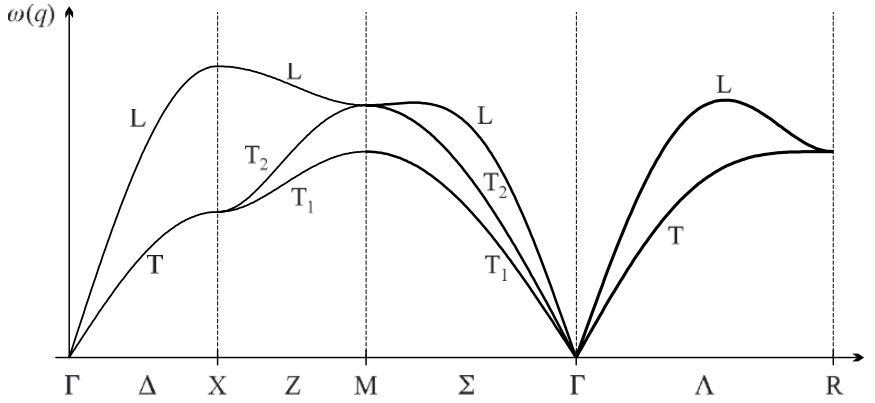


Fig. 11.10. Dispersion relations for the lattice vibrations along the four characteristic directions of the Brillouin zone in a simple cubic crystal with a monatomic basis

sufficiently high symmetry, atoms oscillate in the direction of wave propagation in one branch and in perpendicular directions in the two others. Starting with these excitation branches and continuing them in directions of lower symmetry, one can generally speak of longitudinal or transverse vibrations, although it is not strictly true any more that one eigenvector is parallel and the two others are perpendicular to the propagation direction. The frequency tends to be higher for longitudinal vibrations than for transverse ones. The frequencies of the two transverse vibrations are usually unequal, but degeneracy may occur in directions of high symmetry.

When only first and second neighbors are assumed to interact via central forces, the vibrational spectra for face- and body-centered cubic lattices with a monatomic basis can be calculated along the same lines. The obtained spectra are very similar to the above: they contain three acoustic branches; along directions of sufficiently high symmetry one of them corresponds to longitudinal and two to transverse vibrations. This is in perfect agreement with experimental observations. Figure 11.11 shows the calculated and measured vibrational frequencies in some characteristic directions of the Brillouin zone for a face-centered cubic gold crystal. Measurements were made using neutron scattering techniques.

11.3 The General Description of Lattice Vibrations

In the foregoing we have seen that only longitudinal acoustic vibrations are possible in a linear monatomic chain. In cubic lattices with one atom per primitive cell three acoustic vibrations are possible; in directions of high symmetry one of them is longitudinal (atoms oscillate in the propagation direction) and two are transverse (atoms oscillate in a perpendicular direction). In addition

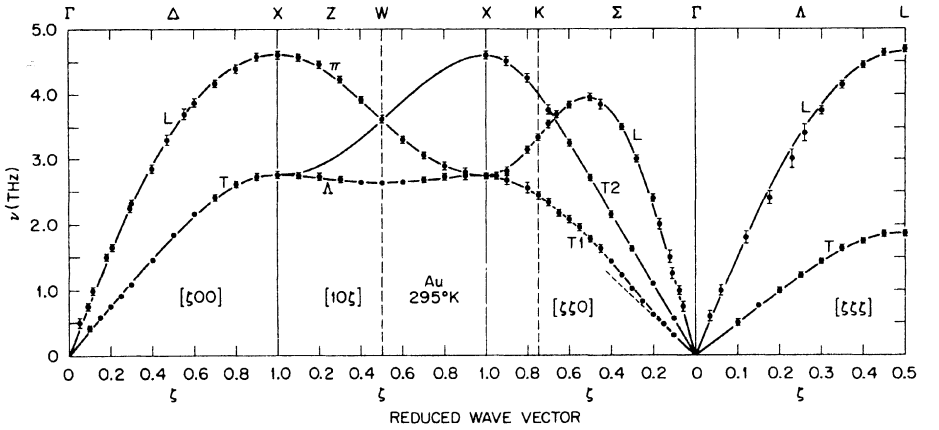


Fig. 11.11. Room-temperature dispersion relation of lattice vibrations for gold in the principal symmetry directions. The reduced wave vector $\zeta = aq/2\pi$. The solid curves were obtained from model calculations with several fitting parameters [J. W. Lynn et al., *Phys. Rev. B* **8**, 3493 (1973)]

to acoustic vibrations, optical vibrations appear in chains with two atoms per primitive cell. These are of finite frequency even in the long-wavelength limit. Below we shall examine the most general case when the three-dimensional crystal contains p atoms per primitive cell.

11.3.1 The Dynamical Matrix and its Eigenvalues

The classical equations of motions were already given in (11.1.28). Just like in the simple examples, their solutions are sought in the form of an expansion in traveling waves – a Fourier series. Since the masses of the p atoms in the primitive cell are not necessarily identical, the explicit separation of a mass-dependent factor proves useful:

$$u_\alpha(m,\mu,t) = \frac{1}{\sqrt{NM_\mu}} \frac{1}{2\pi} \int \mathrm{d}\omega \sum_{\boldsymbol{q}} u_{\mu,\alpha}(\boldsymbol{q}) \mathrm{e}^{\mathrm{i}(\boldsymbol{q}\cdot\boldsymbol{R}_m - \omega t)} \,. \tag{11.3.1}$$

Substituting this form of the displacement into the equation of motion (11.1.28),

$$\omega^2 u_{\mu,\alpha}(\boldsymbol{q}) = \sum_{\nu,\beta} D^{\mu\nu}_{\alpha\beta}(\boldsymbol{q}) u_{\nu,\beta}(\boldsymbol{q}) \,, \tag{11.3.2}$$

where

$$D^{\mu\nu}_{\alpha\beta}(\boldsymbol{q}) = \frac{1}{\sqrt{M_\mu M_\nu}} \sum_n \Phi^{\mu\nu}_{\alpha\beta}(m-n) \mathrm{e}^{-\mathrm{i}\boldsymbol{q}\cdot(\boldsymbol{R}_m - \boldsymbol{R}_n)} \,. \tag{11.3.3}$$

Owing to the translational symmetry of the crystal, $D^{\mu\nu}_{\alpha\beta}(\boldsymbol{q})$ does not depend on the particular choice of the lattice point $\boldsymbol{R}_n$, so we shall use the equivalent expression

$$D^{\mu\nu}_{\alpha\beta}(\boldsymbol{q}) = \frac{1}{\sqrt{M_\mu M_\nu}} \sum_m \Phi^{\mu\nu}_{\alpha\beta}(m) \mathrm{e}^{-\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{R}_m} \,. \tag{11.3.4}$$

This quantity – the Fourier transform of the force constants weighted by the masses – is called the *dynamical matrix*. The system of equations (11.3.2) has nontrivial solutions when

$$\boxed{\det\left(D^{\mu\nu}_{\alpha\beta}(\boldsymbol{q}) - \omega^2 \delta_{\alpha\beta}\delta_{\mu\nu}\right) = 0\,.} \tag{11.3.5}$$

It is important to note that the Fourier coefficients associated with different wave numbers do not mix in this general case, either: for each wave vector $\boldsymbol{q}$ only the $3p$ components – which correspond to the p atoms per primitive cell and the three spatial directions – are mixed. Because of this drastic simplification, the determination of vibrational frequencies is now relatively easy. By arranging the quantities $D^{\mu\nu}_{\alpha\beta}(\boldsymbol{q})$ into a $3p \times 3p$ matrix in the space spanned by the pairs (α, μ) and (β, ν), the previous equation turns out to be the eigenvalue equation of the matrix. We have already encountered a special case, when discussing simple cubic crystals. Consequently, $3p$ solutions are possible for each $\boldsymbol{q}$. We shall now demonstrate that all of them lead to real and positive frequencies.

First we shall show that the dynamical matrix is Hermitian. Substituting the relation (11.1.5) for $\Phi^{\mu\nu}_{\alpha\beta}(m,n)$ into the defining equation (11.3.4), we have

$$D^{\mu\nu}_{\alpha\beta}(\boldsymbol{q}) = \frac{1}{\sqrt{M_\mu M_\nu}} \sum_m \Phi^{\nu\mu}_{\beta\alpha}(-m) \mathrm{e}^{-\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{R}_m} = D^{\nu\mu}_{\beta\alpha}(-\boldsymbol{q})\,. \tag{11.3.6}$$

Since we started with a real potential, the complex phases in the elements of the dynamical matrix must come from the Fourier transform, so

$$D^{\mu\nu}_{\alpha\beta}(\boldsymbol{q}) = D^{\mu\nu\,*}_{\alpha\beta}(-\boldsymbol{q})\,. \tag{11.3.7}$$

Comparison of the two last equations gives

$$D^{\mu\nu}_{\alpha\beta}(\boldsymbol{q}) = D^{\nu\mu\,*}_{\beta\alpha}(\boldsymbol{q})\,. \tag{11.3.8}$$

Strictly speaking, Hermiticity only implies that the $3p$ eigenvalues of the dynamical matrix are all real. In reality, the eigenvalues are not only real but also nonnegative. The reason for this is that the potential is expanded around a minimum, therefore the matrix of force constants is positive definite. This feature is not lost during Fourier transformation, and so the same applies to the dynamical matrix.

Using the notation $\omega^2_\lambda(\boldsymbol{q})$ $(\lambda = 1, 2, \ldots, 3p)$ for the eigenvalues, the eigenvalue equations for the dynamical matrix can also be written in terms of the components of the eigenvectors $e^{(\lambda)}_{\mu,\alpha}$ of the λth eigenvalue:

$$\omega_\lambda^2(\boldsymbol{q})e_{\mu,\alpha}^{(\lambda)}(\boldsymbol{q}) = \sum_{\nu,\beta} D_{\alpha\beta}^{\mu\nu}(\boldsymbol{q})e_{\nu,\beta}^{(\lambda)}(\boldsymbol{q})\,, \tag{11.3.9}$$

or, alternatively,

$$\boxed{\sum_{\nu,\beta}\left[D_{\alpha\beta}^{\mu\nu}(\boldsymbol{q}) - \omega_\lambda^2(\boldsymbol{q})\delta_{\alpha\beta}\delta_{\mu\nu}\right]e_{\nu,\beta}^{(\lambda)}(\boldsymbol{q}) = 0\,.} \tag{11.3.10}$$

This formula is equivalent to (11.3.5).

When eigenvectors are normalized,

$$\sum_{\alpha,\mu} {e_{\mu,\alpha}^{(\lambda)}}^*(\boldsymbol{q})e_{\mu,\alpha}^{(\lambda')}(\boldsymbol{q}) = \delta_{\lambda\lambda'} \tag{11.3.11}$$

follows from their orthogonality, and

$$\sum_{\lambda} {e_{\mu,\alpha}^{(\lambda)}}^*(\boldsymbol{q})e_{\nu,\beta}^{(\lambda)}(\boldsymbol{q}) = \delta_{\alpha\beta}\delta_{\mu\nu} \tag{11.3.12}$$

from their completeness.

11.3.2 Normal Coordinates and Normal Modes

Because of the linearity of the equations of motion, the amplitude of the vibration can be chosen at will, and displacements associated with vibrations of different wave numbers can be freely superposed. By expanding atomic displacements into Fourier series in the spatial variables only, the Fourier component $u_{\mu,\alpha}(\boldsymbol{q},t)$ can now be written as the linear combination of the appropriate components of the unit vectors specifying polarization – i.e., the direction of the displacement of the μth atom relative to the propagation direction determined by $\boldsymbol{q}$. (Hence the name *polarization vector* for $\boldsymbol{e}_\mu^{(\lambda)}(\boldsymbol{q})$.) The time dependence is absorbed into the amplitude $Q_\lambda(\boldsymbol{q},t)$:

$$u_{\mu,\alpha}(\boldsymbol{q},t) = \sum_{\lambda} e_{\mu,\alpha}^{(\lambda)}(\boldsymbol{q})Q_\lambda(\boldsymbol{q},t)\,. \tag{11.3.13}$$

The quantity $Q_\lambda(\boldsymbol{q},t)$ is the *normal coordinate* of the vibrational mode. The instantaneous displacement of the atoms is then written as

$$u_\alpha(m,\mu,t) = \frac{1}{\sqrt{NM_\mu}}\sum_{\boldsymbol{q},\lambda} e_{\mu,\alpha}^{(\lambda)}(\boldsymbol{q})Q_\lambda(\boldsymbol{q},t)\mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{R}_m}\,. \tag{11.3.14}$$

Since normal coordinates are complex, the number of possible values for $\boldsymbol{q}$ being N and that of the label λ being $3p$ apparently implies that the motion of atoms is specified by $6pN$ free parameters. However, as displacements are real, the normal coordinates associated with the wave vectors $\boldsymbol{q}$ and $-\boldsymbol{q}$ are

each others' complex conjugate. Therefore the number of independent normal coordinates is only $3pN$, as it should be.

To prove this, a particularly important consequence of the Hermiticity of the dynamical matrix and the relation (11.3.7) is exploited, namely that the same eigenvalues belong to $\boldsymbol{q}$ and $-\boldsymbol{q}$:

$$\omega_\lambda^2(\boldsymbol{q}) = \omega_\lambda^2(-\boldsymbol{q})\,. \tag{11.3.15}$$

As demonstrated in Chapter 6, this can be regarded as the consequence of the invariance of the equations of motion under time reversal. Since atomic displacements are real, the relation

$$u_{\mu,\alpha}^{(\lambda)}(-\boldsymbol{q},t) = u_{\mu,\alpha}^{(\lambda)\,*}(\boldsymbol{q},t) \tag{11.3.16}$$

must be satisfied, that is,

$$e_{\mu,\alpha}^{(\lambda)}(-\boldsymbol{q})Q_\lambda(-\boldsymbol{q},t) = e_{\mu,\alpha}^{(\lambda)\,*}(\boldsymbol{q})Q_\lambda^*(\boldsymbol{q},t)\,. \tag{11.3.17}$$

Without loss of generality the relation

$$e_{\mu,\alpha}^{(\lambda)\,*}(\boldsymbol{q}) = e_{\mu,\alpha}^{(\lambda)}(-\boldsymbol{q}) \tag{11.3.18}$$

may be imposed, in which case

$$Q_\lambda^*(\boldsymbol{q}) = Q_\lambda(-\boldsymbol{q}) \tag{11.3.19}$$

is indeed satisfied. Note that for crystals with a monatomic basis the vectors $\boldsymbol{e}^{(\lambda)}(\boldsymbol{q})$ can be chosen real.

For future convenience it is useful to represent the decomposition into independent vibrations in yet another way. The classical form (11.1.21) of the potential energy in the harmonic approximation,

$$U_{\text{harm}} = \tfrac{1}{2}\sum_{\substack{m,\mu,\alpha\\ n,\nu,\beta}} \Phi_{\alpha\beta}^{\mu\nu}(m-n)u_\alpha(m,\mu)u_\beta(n,\nu) \tag{11.3.20}$$

is first written in terms of the normal coordinates. Substituting $u_\alpha(m,\mu)$ from (11.3.14) into this formula gives

$$\begin{aligned} U_{\text{harm}} = \tfrac{1}{2}\sum_{\substack{m,\mu,\alpha\\ n,\nu,\beta}} \Phi_{\alpha\beta}^{\mu\nu}(m-n)\frac{1}{\sqrt{NM_\mu}}\sum_{\boldsymbol{q}',\lambda'} e_{\mu,\alpha}^{(\lambda')}(\boldsymbol{q}')Q_{\lambda'}(\boldsymbol{q}')\mathrm{e}^{\mathrm{i}\boldsymbol{q}'\cdot\boldsymbol{R}_m} \\ \times \frac{1}{\sqrt{NM_\nu}}\sum_{\boldsymbol{q},\lambda} e_{\nu,\beta}^{(\lambda)}(\boldsymbol{q})Q_\lambda(\boldsymbol{q})\mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{R}_n}\,. \end{aligned} \tag{11.3.21}$$

To perform the summation over the lattice points, the expressions obtained in (11.3.4) for the dynamical matrix have to be exploited. As $\Phi_{\alpha\beta}^{\mu\nu}$ depends only

on the difference of the lattice vectors, only the terms $\boldsymbol{q}' = -\boldsymbol{q}$ contribute, and so

$$U_{\text{harm}} = \tfrac{1}{2} \sum_{\boldsymbol{q},\lambda,\lambda'} \sum_{\substack{\mu,\alpha \\ \nu,\beta}} D^{\mu\nu}_{\alpha\beta}(\boldsymbol{q}) e^{(\lambda')}_{\mu,\alpha}(-\boldsymbol{q}) Q_{\lambda'}(-\boldsymbol{q}) e^{(\lambda)}_{\nu,\beta}(\boldsymbol{q}) Q_{\lambda}(\boldsymbol{q}) \,. \tag{11.3.22}$$

Making use of the eigenvalue equation (11.3.9),

$$U_{\text{harm}} = \tfrac{1}{2} \sum_{\boldsymbol{q},\lambda,\lambda'} \sum_{\mu,\alpha} \omega^2_{\lambda}(\boldsymbol{q}) e^{(\lambda')}_{\mu,\alpha}(-\boldsymbol{q}) Q_{\lambda'}(-\boldsymbol{q}) e^{(\lambda)}_{\mu,\alpha}(\boldsymbol{q}) Q_{\lambda}(\boldsymbol{q}) \,. \tag{11.3.23}$$

Since displacements are real, and eigenvectors are orthogonal, as expressed in (11.3.17) and (11.3.11), we have

$$U_{\text{harm}} = \tfrac{1}{2} \sum_{\boldsymbol{q},\lambda} \omega^2_{\lambda}(\boldsymbol{q}) Q_{\lambda}(-\boldsymbol{q}) Q_{\lambda}(\boldsymbol{q}) = \tfrac{1}{2} \sum_{\boldsymbol{q},\lambda} \omega^2_{\lambda}(\boldsymbol{q}) Q^*_{\lambda}(\boldsymbol{q}) Q_{\lambda}(\boldsymbol{q}) \,. \tag{11.3.24}$$

The same argument can be used for the expression of the kinetic energy, whereby the formula

$$T_{\text{kin}} = \sum_{m,\mu,\alpha} \tfrac{1}{2} M_{\mu} \dot{u}^2_{\alpha}(m,\mu) \tag{11.3.25}$$

can be rewritten as

$$T_{\text{kin}} = \tfrac{1}{2} \sum_{\boldsymbol{q},\lambda} \dot{Q}_{\lambda}(-\boldsymbol{q}) \dot{Q}_{\lambda}(\boldsymbol{q}) = \tfrac{1}{2} \sum_{\boldsymbol{q},\lambda} \dot{Q}^*_{\lambda}(\boldsymbol{q}) \dot{Q}_{\lambda}(\boldsymbol{q}) \,. \tag{11.3.26}$$

Once the Lagrangian $\mathcal{L} = T_{\text{kin}} - U_{\text{harm}}$ is known, the equations of motion for the normal coordinates are readily established from Lagrange's equation[2]

$$\frac{\partial}{\partial t} \frac{\partial \mathcal{L}}{\partial \dot{Q}_{\lambda}(-\boldsymbol{q})} - \frac{\partial \mathcal{L}}{\partial Q_{\lambda}(-\boldsymbol{q})} = 0 \,. \tag{11.3.27}$$

Then

$$\ddot{Q}_{\lambda}(\boldsymbol{q}) = -\omega^2_{\lambda}(\boldsymbol{q}) Q_{\lambda}(\boldsymbol{q}) \tag{11.3.28}$$

for each normal coordinate – that is, using normal coordinates, the vibrations of the system are automatically separated into those of independent harmonic oscillators. These are called the *normal modes of vibration*. Since λ can take $3p$ different values in (11.3.28), and the number of possible values for $\boldsymbol{q}$ is N, the three-dimensional coupled vibrations of the crystal (made up of pN atoms) can be described with $3pN$ independent oscillators.

To prepare the quantum mechanical discussion of the next chapter, it is useful to write down the classical Hamiltonian using normal coordinates.

[2] More precisely: Lagrange's equation of the second kind. Since they are of the same form as Euler's equation in variational calculus, the name *Euler–Lagrange equation* is also used.

Following the steps of the Hamiltonian formulation in classical mechanics, we first introduce the canonical conjugate to the normal coordinate $Q_\lambda(\boldsymbol{q})$, which is the derivative of the Lagrangian with respect to $\dot{Q}_\lambda(\boldsymbol{q})$:

$$P_\lambda(\boldsymbol{q}) = \frac{\partial \mathcal{L}}{\partial \dot{Q}_\lambda(\boldsymbol{q})} = \dot{Q}^*_\lambda(\boldsymbol{q})\,, \tag{11.3.29}$$

and then the Hamiltonian

$$\begin{aligned} \mathcal{H}_{\text{harm}} &= \sum_{\boldsymbol{q},\lambda} P_\lambda(\boldsymbol{q})\dot{Q}_\lambda(\boldsymbol{q}) - \mathcal{L} \\ &= \tfrac{1}{2}\sum_{\boldsymbol{q},\lambda} \left\{ P^*_\lambda(\boldsymbol{q})P_\lambda(\boldsymbol{q}) + \omega^2_\lambda(\boldsymbol{q})Q^*_\lambda(\boldsymbol{q})Q_\lambda(\boldsymbol{q}) \right\} \\ &= \tfrac{1}{2}\sum_{\boldsymbol{q},\lambda} \left\{ |P_\lambda(\boldsymbol{q})|^2 + \omega^2_\lambda(\boldsymbol{q})|Q_\lambda(\boldsymbol{q})|^2 \right\}. \end{aligned} \tag{11.3.30}$$

The full Hamiltonian is the sum of the Hamiltonians for individual normal modes. The equations of motion obtained in the Hamiltonian formulation,

$$\begin{aligned} \dot{Q}_\lambda(\boldsymbol{q}) &= \frac{\partial \mathcal{H}}{\partial P_\lambda(\boldsymbol{q})} = P^*_\lambda(\boldsymbol{q})\,, \\ \dot{P}_\lambda(\boldsymbol{q}) &= -\frac{\partial \mathcal{H}}{\partial Q_\lambda(\boldsymbol{q})} = -\omega^2_\lambda(\boldsymbol{q})Q^*_\lambda(\boldsymbol{q}) \end{aligned} \tag{11.3.31}$$

are equivalent to (11.3.28).

11.3.3 Acoustic and Optical Vibrations

In the previous subsections we saw that if the basis contains p atoms, there exist $3p$ vibrational frequencies for each vector $\boldsymbol{q}$ of the Brillouin zone. In the one-dimensional case we also saw that some of them correspond to acoustic vibrations, while others to optical ones. As a generalization we shall demonstrate that among the branches formed by the $3p$ vibrational frequencies three can always be called acoustic, since they start from $\omega = 0$ at $\boldsymbol{q} = 0$, and the others optical, as their frequency does not vanish at $\boldsymbol{q} = 0$.

Acoustic Vibrations

A branch of vibrations is called acoustic if it has a vanishing frequency at $\boldsymbol{q} = 0$, the center of the Brillouin zone. We shall now prove that there always exist three such branches. When the eigenvalue equation (11.3.9) is examined at $\boldsymbol{q} = 0$, and the value of the dynamical matrix at $\boldsymbol{q} = 0$ is taken from (11.3.4), we have

$$\omega^2_\lambda(0)e^{(\lambda)}_{\mu,\alpha}(0) = \sum_{\nu,\beta} D^{\mu\nu}_{\alpha\beta}(0)e^{(\lambda)}_{\nu,\beta}(0) = \sum_{m,\nu,\beta} \frac{1}{\sqrt{M_\mu M_\nu}} \Phi^{\mu\nu}_{\alpha\beta}(m)e^{(\lambda)}_{\nu,\beta}(0)\,. \tag{11.3.32}$$

If the quantity $e^{(\lambda)}_{\nu,\beta}(0)/\sqrt{M_\nu}$ is constant and independent of ν for the three Cartesian coordinates β, that is

$$e^{(\lambda)}_{\nu,\beta} = C\sqrt{M_\nu}e^{(\lambda)}_{\beta}\,, \tag{11.3.33}$$

where the normalization factor

$$C = \left(\sum_{\mu=1}^{p} M_\mu\right)^{-1/2} \tag{11.3.34}$$

ensures that the orthogonality relation of the polarization vectors is satisfied by vectors $\boldsymbol{e}^{(\lambda)}$ of unit length, then (11.1.8) implies the vanishing of the right-hand side of (11.3.32), so the frequency on the left-hand side must also be zero.

In vector form this condition reads

$$\boldsymbol{e}^{(\lambda)}_{\nu} = \sqrt{M_\nu}\boldsymbol{e}^{(\lambda)}. \tag{11.3.35}$$

Since three mutually perpendicular vectors can be chosen for the $\boldsymbol{e}^{(\lambda)}$, there are three solutions with vanishing frequency at $\boldsymbol{q} = 0$. For such vibrations the atomic displacements

$$u_\alpha(m,\mu) = \frac{1}{\sqrt{NM_\mu}}e^{(\lambda)}_{\mu,\alpha}Q_\lambda(0) = \frac{1}{\sqrt{N}}e^{(\lambda)}_{\alpha}Q_\lambda(0) \tag{11.3.36}$$

are independent of the label μ; in other words, all the atoms of the primitive cell oscillate in phase, with equal vibrational amplitudes. Owing to continuity, there are three similar vibrations for small values of q. Atoms in the primitive cell vibrate almost in phase with almost equal amplitudes – just as in a sound wave (long-wavelength elastic wave) propagating in a solid. The dispersion relation of these vibrations give the three acoustic branches.

Optical Vibrations

In addition to the three acoustic branches, there are $3p-3$ further vibrational branches in which frequencies generally start at some finite value. These are the branches of optical modes. To understand the character of such vibrations, consider a crystal with a diatomic basis. It follows from (11.3.35) that for vibrations of polarization λ in the acoustic branch

$$\frac{\boldsymbol{e}^{(\lambda)}_1}{\sqrt{M_1}} = \frac{\boldsymbol{e}^{(\lambda)}_2}{\sqrt{M_2}} \tag{11.3.37}$$

at the center of the Brillouin zone. Whichever optical branch with polarization λ' is chosen,

$$\boldsymbol{e}^{(\lambda)}_1 \cdot \boldsymbol{e}^{(\lambda')}_1 + \boldsymbol{e}^{(\lambda)}_2 \cdot \boldsymbol{e}^{(\lambda')}_2 = 0 \tag{11.3.38}$$

because of the orthogonality of eigenvectors. From this pair of equations

$$\boldsymbol{e}_1^{(\lambda)} \cdot \left[\boldsymbol{e}_1^{(\lambda')} + \sqrt{\frac{M_2}{M_1}} \boldsymbol{e}_2^{(\lambda')} \right] = 0 \,. \tag{11.3.39}$$

This condition must hold for each of the three acoustic branches, consequently

$$\sqrt{M_1} \boldsymbol{e}_1^{(\lambda')} + \sqrt{M_2} \boldsymbol{e}_2^{(\lambda')} = 0 \tag{11.3.40}$$

for each optical branch. Equation (11.3.14) permits us to set up a relation between the polarization vector and the atomic displacements due to vibrations of wave number $\boldsymbol{q} = 0$. The above equation then implies that for the displacements due to optical vibrations

$$M_1 \boldsymbol{u}_1(m, 1) + M_2 \boldsymbol{u}_2(m, 2) = 0 \,, \tag{11.3.41}$$

that is, in long-wavelength optical vibrations atoms in the primitive cell move in opposite directions in such a way that the center of mass remains stationary. Therefore these vibrations may be regarded as internal oscillations of the system of particles inside a primitive cell.

Among optical vibrations one may distinguish longitudinal and transverse modes as well – at least in directions of high symmetry. Depending on the symmetry in question, frequencies may become degenerate. Figure 11.12 shows the spectra of lattice vibrations in some characteristic directions of the Brillouin zone for crystalline silicon, which contains two atoms per primitive cell in a diamond structure.

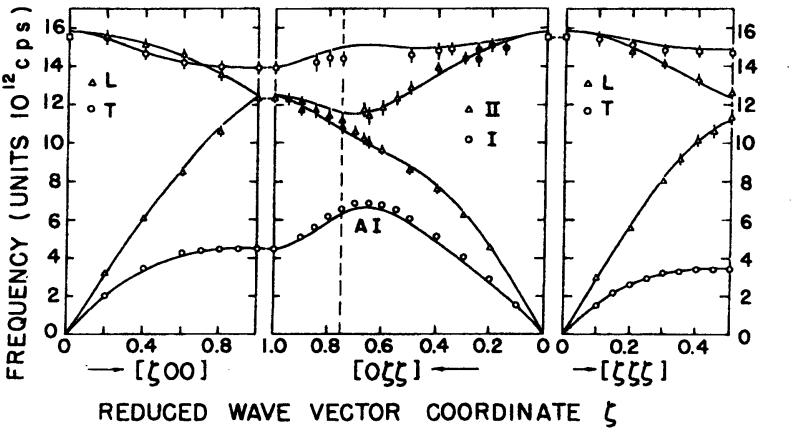


Fig. 11.12. Dispersion relations in some characteristic directions of the Brillouin zone for crystalline silicon, which contains two atoms per primitive cell. The solid line is a calculated curve fitted to experimental data. [G. Dolling, in *Inelastic Scattering of Neutrons in Solids and Liquids*, IAEA, Vienna, 1963, Vol. II., p. 37]

Besides experimental results, theoretical calculations are also shown in the figure. Assuming central forces, a relatively large number of force constants have to be introduced among distant neighbors to ensure a good fit. The reason for this is that approximating highly directional covalent bonds by springs is difficult to justify. Another drawback of the picture of springs connecting atoms is that the polarization of electron clouds due to the motion of ions cannot be appropriately taken into account in calculations of the vibrational spectrum. The proper treatment of such effects requires much more sophisticated methods that cannot be discussed here.

11.4 Lattice Vibrations in the Long-Wavelength Limit

In the previous section we saw that in long-wavelength acoustic modes the atoms in a primitive cell oscillate in phase and with equal amplitudes. The displacement of the atoms in the mth primitive cell – characterized by the lattice vector $\boldsymbol{R}_m$ – is independent of the label μ in the first approximation, therefore the internal structure of the primitive cell (the basis) is irrelevant from the viewpoint of vibrations. Moreover, the common displacement $\boldsymbol{u}(\boldsymbol{R}_m)$ varies little between adjacent cells. For this reason, a slowly varying continuous function $\boldsymbol{u}(\boldsymbol{r})$ may be introduced, whose value in the lattice point $\boldsymbol{R}_m$ is the displacement of the atoms of the primitive cell. This quantity specifies the displacement with respect to the equilibrium position in an arbitrary point $\boldsymbol{r}$ of the solid when atomic details are ignored – i.e., when matter is considered to be continuous. In this approximation the solid can be regarded as a continuous elastic medium characterized by a handful of elastic constants, in which classical elastic waves may propagate. Below we shall examine the relationship between these waves and atomic oscillations.

This approach cannot be applied to optical vibrations. Instead, a simple relation will be established between the frequency of optical vibrations and dielectric polarizability in ionic crystals.

11.4.1 Acoustic Vibrations as Elastic Waves

The starting point for our investigations into the long-wavelength limit of acoustic vibrations is (11.1.10), a formula of general validity when interatomic interactions are treated in the harmonic approximation. As atoms in a primitive cell move together – i.e., $\boldsymbol{u}(m,\mu)$ is essentially the same for all atoms in the primitive cell at $\boldsymbol{R}_m$ –, and $\Phi^{\mu\nu}_{\alpha\beta}(m,n)$ depends only on the difference of vectors $\boldsymbol{R}_m$ and $\boldsymbol{R}_n$, the harmonic term can be written in a simple form, similar to that in crystals with a monatomic basis:

$$U_{\text{harm}} = -\tfrac{1}{4}\sum_{\substack{m,n\\ \alpha,\beta}} [u_\alpha(m) - u_\alpha(n)]\, \Phi_{\alpha\beta}(m-n)\, [u_\beta(m) - u_\beta(n)]\ , \qquad (11.4.1)$$

where $\Phi_{\alpha\beta}(m-n)$ is an effective force constant between the primitive cells at $\boldsymbol{R}_m$ and $\boldsymbol{R}_n$.

If atomic displacements vary little over the range of the force constants then the difference in the displacement of the atoms in the primitive cells at $\boldsymbol{R}_m$ and $\boldsymbol{R}_n$ is well approximated by the first derivative of the continuous displacement field $\boldsymbol{u}(\boldsymbol{r})$:

$$\boldsymbol{u}(m) = \boldsymbol{u}(n) + ((\boldsymbol{r}_m - \boldsymbol{r}_n) \cdot \boldsymbol{\nabla}_{\boldsymbol{r}})\, \boldsymbol{u}(\boldsymbol{r})|_{\boldsymbol{r}=\boldsymbol{R}_n} . \tag{11.4.2}$$

Using this expansion, the harmonic term of the interaction is

$$U_{\text{harm}} = \tfrac{1}{2} \sum_{\substack{n,\alpha,\beta \\ \sigma,\tau}} \left(\frac{\partial}{\partial r_\sigma} u_\alpha(\boldsymbol{r}) \right)_{\boldsymbol{r}=\boldsymbol{R}_n} \left(\frac{\partial}{\partial r_\tau} u_\beta(\boldsymbol{r}) \right)_{\boldsymbol{r}=\boldsymbol{R}_n} E^{\sigma\tau}_{\alpha\beta}(\boldsymbol{R}_n) , \tag{11.4.3}$$

where

$$E^{\sigma\tau}_{\alpha\beta}(\boldsymbol{R}_n) = -\tfrac{1}{2} \sum_m [\boldsymbol{R}_m - \boldsymbol{R}_n]_\sigma \Phi_{\alpha\beta}(m-n) [\boldsymbol{R}_m - \boldsymbol{R}_n]_\tau . \tag{11.4.4}$$

In a homogeneous crystal this quantity is independent of the position vector $\boldsymbol{R}_n$. Since the function $\boldsymbol{u}(\boldsymbol{r})$ varies slowly, the above sum is well approximated by the integral

$$U_{\text{harm}} = \frac{1}{2v} \sum_{\substack{\sigma\tau \\ \alpha\beta}} \int \mathrm{d}\boldsymbol{r} \left(\frac{\partial}{\partial r_\sigma} u_\alpha(\boldsymbol{r}) \right) \left(\frac{\partial}{\partial r_\tau} u_\beta(\boldsymbol{r}) \right) E^{\sigma\tau}_{\alpha\beta} , \tag{11.4.5}$$

where v is the volume of the primitive cell. This expression gives the energy of the elastically deformed crystal in terms of a four-index tensor $E^{\sigma\tau}_{\alpha\beta}$, whose elements could be taken as the elastic constants. Due to some symmetry properties, $E^{\sigma\tau}_{\alpha\beta}$ does not have 81 different components even in the most general case: the number of independent elastic constants is much smaller.

To prove this and to find the appropriate constants, we first make use of the property that the antisymmetric combination

$$\tfrac{1}{2} \left(\frac{\partial}{\partial r_\beta} u_\alpha - \frac{\partial}{\partial r_\alpha} u_\beta \right) \tag{11.4.6}$$

corresponds to a rigid rotation of the solid, and so it gives a vanishing contribution to the energy. Therefore only the symmetric combination

$$\varepsilon_{\alpha\beta} = \tfrac{1}{2} \left(\frac{\partial}{\partial r_\beta} u_\alpha + \frac{\partial}{\partial r_\alpha} u_\beta \right) \tag{11.4.7}$$

may appear in the energy expression. Among the elements $\varepsilon_{\alpha\beta}$ of the *strain tensor*, diagonal ones (ε_{xx}, ε_{yy}, and ε_{zz}) are related to the relative change in length along the coordinate axes, while off-diagonal elements describe shear strains. Elastic energy is a homogeneous second-order expression of these

quantities. Which combinations appear is determined by the symmetries of the system, since the energy expression must be invariant under all symmetry operations that take the system into itself. This requirement imposes severe restrictions on the number of independent elastic constants.

As an example, consider a homogeneous and isotropic (therefore noncrystalline) material. In this case only those second-order expressions may appear that are invariant under arbitrary rotations of the coordinate system. By its definition, the strain tensor is symmetric: $\varepsilon_{\alpha\beta} = \varepsilon_{\beta\alpha}$. Such a tensor has one linear invariant, the sum of its diagonal elements (its trace):

$$I_1 = \varepsilon_{xx} + \varepsilon_{yy} + \varepsilon_{zz}\,, \tag{11.4.8}$$

which is equal to the relative variation of the volume. Besides I_1^2, there is another (independent) quadratic invariant,

$$I_2 = \varepsilon_{xx}\varepsilon_{yy} + \varepsilon_{yy}\varepsilon_{zz} + \varepsilon_{zz}\varepsilon_{xx} - \tfrac{1}{2}\left(\varepsilon_{xy}^2 + \varepsilon_{yx}^2 + \varepsilon_{yz}^2 + \varepsilon_{zy}^2 + \varepsilon_{zx}^2 + \varepsilon_{xz}^2\right). \tag{11.4.9}$$

Using the combination $I_1^2 - 2I_2$ instead, only two parameters are needed to write down the elastic energy in the isotropic case:

$$\begin{aligned} U_{\text{harm}} = {} & \frac{\lambda}{2}\left[\varepsilon_{xx} + \varepsilon_{yy} + \varepsilon_{zz}\right]^2 \\ & + \mu\left[\varepsilon_{xx}^2 + \varepsilon_{yy}^2 + \varepsilon_{zz}^2 + \varepsilon_{xy}^2 + \varepsilon_{yx}^2 + \varepsilon_{yz}^2 + \varepsilon_{zy}^2 + \varepsilon_{zx}^2 + \varepsilon_{xz}^2\right] \end{aligned} \tag{11.4.10}$$

where λ and μ are the *Lamé constants*:[3] μ is the usual shear modulus G, while λ is related to the compressibility κ and the bulk modulus K via

$$\frac{1}{\kappa} = K = \lambda + \frac{2}{3}\mu\,. \tag{11.4.11}$$

In elasticity, stresses generated by elastic strains are characterized by the symmetric *stress tensor* $\sigma_{\alpha\beta}$, and the force per unit volume of the deformed body is given by the derivative

$$F_\alpha = \sum_\beta \frac{\partial \sigma_{\alpha\beta}}{\partial r_\beta}\,. \tag{11.4.12}$$

The vibrations of an elastic medium of density ρ are therefore governed by the equation of motion

$$\rho\,\ddot{u}_\alpha = \sum_\beta \frac{\partial \sigma_{\alpha\beta}}{\partial r_\beta}\,. \tag{11.4.13}$$

To put this to use, another relation is needed: a generalization of Hooke's law[4] that establishes the connection between the stress tensor and the strain tensor. This can be written as

[3] G. Lamé, 1852.
[4] R. Hooke, 1676.

$$\sigma_{\alpha\beta} = \delta_{\alpha\beta}\lambda\,[\varepsilon_{xx} + \varepsilon_{yy} + \varepsilon_{zz}] + 2\mu\,\varepsilon_{\alpha\beta}\,, \tag{11.4.14}$$

derived from the relation

$$\sigma_{\alpha\beta} = \frac{\partial U_{\mathrm{harm}}}{\partial \varepsilon_{\alpha\beta}}\,. \tag{11.4.15}$$

We are now in the position to determine the propagation velocity of waves in an isotropic elastic medium. Writing out the equation of motion for the component u_x in full detail,

$$\begin{aligned}
\rho\,\ddot{u}_x &= \frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{xy}}{\partial y} + \frac{\partial \sigma_{xz}}{\partial z} \\
&= \lambda \frac{\partial}{\partial x}\left(\frac{\partial u_x}{\partial x} + \frac{\partial u_y}{\partial y} + \frac{\partial u_z}{\partial z}\right) + 2\mu \frac{\partial^2 u_x}{\partial x^2} \\
&\quad + \mu\left(\frac{\partial^2 u_x}{\partial y^2} + \frac{\partial^2 u_x}{\partial z^2} + \frac{\partial^2 u_y}{\partial x \partial y} + \frac{\partial^2 u_z}{\partial x \partial z}\right) \\
&= (\lambda + \mu)\frac{\partial}{\partial x}\left(\frac{\partial u_x}{\partial x} + \frac{\partial u_y}{\partial y} + \frac{\partial u_z}{\partial z}\right) + \mu\,\nabla^2 u_x\,.
\end{aligned} \tag{11.4.16}$$

Similar equations apply to other components. The three components can be written concisely in the vector form

$$\rho\,\ddot{\boldsymbol{u}} = (\lambda + \mu)\ \mathrm{grad}\,\mathrm{div}\,\boldsymbol{u} + \mu\,\nabla^2 \boldsymbol{u}\,. \tag{11.4.17}$$

Solutions that are periodic in space and time are sought in the form

$$\boldsymbol{u}(\boldsymbol{r}) = \boldsymbol{e}\,\mathrm{e}^{\mathrm{i}(\boldsymbol{q}\cdot\boldsymbol{r} - \omega t)}\,, \tag{11.4.18}$$

which leads to a relationship between $\boldsymbol{e}$, $\boldsymbol{q}$ and ω:

$$\rho\,\omega^2 \boldsymbol{e} = (\lambda + \mu)\,\boldsymbol{q}(\boldsymbol{q}\cdot\boldsymbol{e}) + \mu\,\boldsymbol{q}^2 \boldsymbol{e}\,. \tag{11.4.19}$$

Decomposing the vector $\boldsymbol{e}$ that represents the displacement amplitude into components parallel to the vector $\boldsymbol{q}$ specifying the direction of propagation (longitudinal wave) and perpendicular to it (transverse wave),

$$\begin{aligned}
\rho\,\omega^2 \boldsymbol{e}_{\mathrm{L}} &= (\lambda + \mu)\,q^2 \boldsymbol{e}_{\mathrm{L}} + \mu\,q^2 \boldsymbol{e}_{\mathrm{L}}\,, \\
\rho\,\omega^2 \boldsymbol{e}_{\mathrm{T}} &= \mu\,q^2 \boldsymbol{e}_{\mathrm{T}}\,.
\end{aligned} \tag{11.4.20}$$

It is readily seen that the dispersion relation is linear for longitudinal (compressional) and transverse (torsional) waves alike:

$$\omega_{\mathrm{L}} = c_{\mathrm{L}}\,q\,, \qquad \omega_{\mathrm{T}} = c_{\mathrm{T}}\,q\,, \tag{11.4.21}$$

however, their propagation velocities are different:

$$c_{\mathrm{L}} = \sqrt{\frac{\lambda + 2\mu}{\rho}}\,, \qquad c_{\mathrm{T}} = \sqrt{\frac{\mu}{\rho}}\,. \tag{11.4.22}$$

Table 11.1. Propagation velocities of long-wavelength longitudinal and transverse waves in some metals at room temperature

Element	c_{L} $(\mathrm{m\,s^{-1}})$	c_{T} $(\mathrm{m\,s^{-1}})$	Element	c_{L} $(\mathrm{m\,s^{-1}})$	c_{T} $(\mathrm{m\,s^{-1}})$
Ag	3640	1690	Mg	5700	3170
Al	6360	3130	Ni	5810	3080
Au	3280	1190	Pb	2050	710
Be	12720	8330	Pt	4080	1690
Cr	6850	3980	Sn	3300	1650
Cu	4760	2300	Ti	6260	2920
Fe	5920	3220	V	6000	2780

In crystals, the velocity of longitudinal waves (sound) is usually around several thousand m/s. Transverse waves usually propagate more slowly, roughly at half their speed, mostly in the range 700 to 3500 m/s. Values measured in some metals are listed in Table 11.1.

A vibration may be considered of "long wavelength" if its wavelength is at least one order of magnitude larger than the interatomic spacing. On such length scales solids can be described to a good approximation as elastic continua. The discrete character of the lattice becomes important only for vibrations of shorter wavelengths. In terms of frequencies, the interpretation of acoustic lattice vibrations as elastic waves is justified for frequencies up to the order of 10^{11} Hz.

11.4.2 Elastic Constants of Crystalline Materials

The model of elastic continua may also be applied to the discussion of the elastic properties of crystals. However, instead of (11.4.14) – Hooke's law, establishing a linear relationship between stress and strain for isotropic systems – an even more general relationship is required between the stress tensor and the strain tensor:

$$\sigma_{\alpha\beta} = \sum_{\gamma\delta} C_{\alpha\beta,\gamma\delta}\varepsilon_{\gamma\delta}\,. \tag{11.4.23}$$

Instead of (11.4.5), the elastic energy is now expressed as

$$U_{\mathrm{harm}} = \tfrac{1}{2}\sum_{\alpha\beta\gamma\delta} \varepsilon_{\alpha\beta} C_{\alpha\beta,\gamma\delta}\varepsilon_{\gamma\delta}\,. \tag{11.4.24}$$

Elastic properties are thus characterized by a four-index tensor. However, since both the stress and strain tensors are symmetric,

$$C_{\alpha\beta,\gamma\delta} = C_{\beta\alpha,\gamma\delta} = C_{\alpha\beta,\delta\gamma} = C_{\gamma\delta,\alpha\beta}\,. \tag{11.4.25}$$

Depending on the symmetries of the crystal, further relationships may exist among the tensor elements.

Instead of the four-index elastic constants, the *Voigt elastic constants*[5] are used in the literature. These are obtained by using the notation

$$\begin{aligned} \varepsilon_{xx} &= \varepsilon_1\,, & \varepsilon_{yy} &= \varepsilon_2\,, & \varepsilon_{zz} &= \varepsilon_3\,, \\ 2\varepsilon_{yz} &= \varepsilon_4\,, & 2\varepsilon_{zx} &= \varepsilon_5\,, & 2\varepsilon_{xy} &= \varepsilon_6 \end{aligned} \tag{11.4.26}$$

for the elements of the symmetrical strain tensor. Using these quantities, the homogeneous quadratic expression for the elastic energy density reads

$$U_{\text{harm}} = \tfrac{1}{2} \sum_{i,j=1}^{6} c_{ij}\varepsilon_i\varepsilon_j\,, \tag{11.4.27}$$

where the coefficients c_{ij} are the Voigt elastic constants. A similar notation is introduced for the elements of the stress tensor:

$$\begin{aligned} \sigma_{xx} &= \sigma_1\,, & \sigma_{yy} &= \sigma_2\,, & \sigma_{zz} &= \sigma_3\,, \\ \sigma_{yz} &= \sigma_4\,, & \sigma_{zx} &= \sigma_5\,, & \sigma_{xy} &= \sigma_6\,. \end{aligned} \tag{11.4.28}$$

The generalized Hooke's law, which establishes the relation between the stress tensor and the strain tensor, now takes the form

$$\sigma_i = \frac{\partial U_{\text{harm}}}{\partial \varepsilon_i} = \sum_{j=1}^{6} c_{ij}\varepsilon_j\,. \tag{11.4.29}$$

By its definition, the matrix of elastic constants is also symmetric, therefore in the most general case the elastic properties of triclinic crystals are characterized by 21 independent constants.

In crystals of higher symmetry the number of independent elastic constants is lower. For example, when the z-axis is chosen as the twofold rotation axis in the monoclinic crystal system, elastic energy must be invariant under the transformation $x \to -x$, $y \to -y$, $z \to z$. In such rotations ε_4 and ε_5 change sign but the four other elements of the strain tensor do not. Since the elastic energy has to remain invariant,

$$c_{14} = c_{24} = c_{34} = c_{15} = c_{25} = c_{35} = c_{46} = c_{56} = 0\,. \tag{11.4.30}$$

Therefore in monoclinic crystals there are 13 independent elastic constants.

In the orthorhombic crystal system elastic energy must be invariant under 180° rotations around each of the three axes. This requires

$$c_{16} = c_{26} = c_{36} = c_{45} = 0\,, \tag{11.4.31}$$

so elastic behavior is characterized by 9 elastic constants.

[5] W. Voigt, 1910.

In tetragonal crystals rotation through 90° around the z-axis appears as a new symmetry. Invariance of the elastic energy under the transformation $x \to y$, $y \to -x$, $z \to z$ requires that

$$c_{11} = c_{22}\,, \qquad c_{13} = c_{23}\,, \qquad c_{44} = c_{55}\,. \tag{11.4.32}$$

This reduces the number of independent elastic constants to 6. The expression for the elastic energy is then

$$\begin{aligned} U_{\text{harm}} = {} & \tfrac{1}{2}c_{11}\left[\varepsilon_{xx}^2 + \varepsilon_{yy}^2\right] + \tfrac{1}{2}c_{33}\varepsilon_{zz}^2 + c_{12}\varepsilon_{xx}\varepsilon_{yy} \\ & + c_{13}\left[\varepsilon_{xx}\varepsilon_{zz} + \varepsilon_{yy}\varepsilon_{zz}\right] + 2c_{44}\left[\varepsilon_{yz}^2 + \varepsilon_{zx}^2\right] + 2c_{66}\varepsilon_{xy}^2\,. \end{aligned} \tag{11.4.33}$$

In cubic crystals rotations around the x- and y-axes through 90° ($x \to x$, $y \to z$, $z \to -y$, and $y \to y$, $x \to z$, $z \to -x$) are also symmetries. These imply further restrictions:

$$c_{33} = c_{11}\,, \qquad c_{13} = c_{12}\,, \qquad c_{66} = c_{44}\,. \tag{11.4.34}$$

In terms of the three remaining elastic constants,

$$\begin{aligned} U_{\text{harm}} = {} & \tfrac{1}{2}c_{11}\left[\varepsilon_{xx}^2 + \varepsilon_{yy}^2 + \varepsilon_{zz}^2\right] + c_{12}\left[\varepsilon_{xx}\varepsilon_{yy} + \varepsilon_{xx}\varepsilon_{zz} + \varepsilon_{yy}\varepsilon_{zz}\right] \\ & + 2c_{44}\left[\varepsilon_{xy}^2 + \varepsilon_{yz}^2 + \varepsilon_{zx}^2\right]. \end{aligned} \tag{11.4.35}$$

It can be shown in much the same manner that the number of independent elastic constants is 6 for the rhombohedral and 5 for the hexagonal crystal system.

Further connections called *Cauchy relations* exist among the elastic constants when interatomic forces are all central. For materials of cubic crystal structure

$$c_{12} = c_{44} \tag{11.4.36}$$

in this case. Elastic constants determined at room temperature are shown in Table 11.2 for some materials with cubic crystal structure. Values measured at low temperatures are somewhat higher but their deviation from the listed ones does not exceed ten to twenty percent.

As the table shows, the Cauchy relation is almost perfectly satisfied in ionic and covalent crystals, while significant deviations are observed in the majority of metals. This indicates the importance of taking noncentral forces into account in the theoretical determination of elastic constants.

It should be noted that instead of the Voigt constants elastic behavior is often characterized by four elastic moduli – Young's modulus or the tensile modulus (also known as the modulus of elasticity) E, the bulk modulus or compression modulus K (the inverse of the compressibility κ), Poisson's ratio ν related to lateral contraction, and the shear modulus or modulus of rigidity G. The relationships among them and their connections to the Lamé and

Table 11.2. Voigt elastic constants (in units of GPa) for some materials of cubic crystal structure

Material	c_{11}	c_{12}	c_{44}	Material	c_{11}	c_{12}	c_{44}
Na	7.59	6.33	4.30	C	1079	124.5	578
K	3.69	3.18	1.90	Si	165.6	63.9	79.6
Rb	2.96	2.44	1.60	Ge	129	48.3	67.1
Cr	348	67	100.0	Pb	48.8	41.4	14.8
Mo	465	163	109	NaCl	49.47	12.88	12.87
W	523	203	160	KCl	40.69	7.11	6.31
Fe	230	135	117	CsCl	36.44	8.82	8.04
Ir	580	242	256	NaBr	39.7	10.01	9.98
Ni	247	153	122	GaAs	118.8	53.7	59.4
Pt	347	251	76.5	ZnS	104.6	65.3	46.1
Cu	169	122	75.3	Fe_3O_4	273	106	97.1
Ag	122	92	45.5	TiC	500	113	175
Au	191	162	42.2	MgO	297.1	95.4	156.1
Al	108	62	28.3	$MgAl_2O_4$	298	154	158

Voigt constants are presented in most textbooks on elasticity. For example, in isotropic materials the elastic moduli can be simply expressed in terms of the Lamé constants:

$$K = \lambda + \tfrac{2}{3}\mu, \quad G = \mu, \quad \nu = \frac{\lambda}{2(\lambda+\mu)}, \quad E = \frac{\mu(3\lambda+2\mu)}{\lambda+\mu}, \tag{11.4.37}$$

or conversely, the Lamé constants can be given in terms of two independent moduli, E and ν or K and ν as

$$\begin{aligned} \lambda &= \frac{\nu E}{(1+\nu)(1-2\nu)} = 3K\frac{\nu}{1+\nu}, \\ \mu &= \frac{E}{2(1+\nu)} = 3K\frac{1-2\nu}{2(1+\nu)}. \end{aligned} \tag{11.4.38}$$

For cubic crystals

$$K = \tfrac{1}{3}(c_{11}+2c_{12}), \qquad G = \tfrac{1}{2}(c_{11}-c_{12}), \qquad \nu = \frac{c_{12}}{c_{11}+c_{12}}, \tag{11.4.39}$$

and, depending on the direction relative to the crystallographic axes, Young's modulus is

$$\begin{aligned} E_{[100]} &= \frac{(c_{11}-c_{12})(c_{11}+2c_{12})}{c_{11}+c_{12}}, \\ E_{[110]} &= \frac{4c_{44}(c_{11}-c_{12})(c_{11}+2c_{12})}{2c_{11}c_{44}+(c_{11}-c_{12})(c_{11}+2c_{12})}, \\ E_{[111]} &= \frac{3c_{44}(c_{11}+2c_{12})}{c_{11}+2c_{12}+c_{44}}. \end{aligned} \tag{11.4.40}$$

As we have already seen, only two independent elastic constants remain in the isotropic case. In terms of these the tensor $C_{\alpha\beta,\gamma\delta}$ is written as

$$C_{\alpha\beta,\gamma\delta} = \lambda\delta_{\alpha\beta}\delta_{\gamma\delta} + \mu\left(\delta_{\alpha\gamma}\delta_{\beta\delta} + \delta_{\alpha\delta}\delta_{\beta\gamma}\right) . \tag{11.4.41}$$

Using this expression in the elastic energy formula, and comparing it with the corresponding expression for cubic crystals written in terms of the three independent Voigt constants, the following relations are obtained:

$$\lambda = c_{12} , \qquad \mu = c_{44} , \tag{11.4.42}$$

and $\lambda + 2\mu = c_{11}$. This implies the relationship

$$s \equiv \frac{c_{11} - c_{12}}{2c_{44}} = 1 \tag{11.4.43}$$

for isotropic materials. Therefore deviations from isotropy may be characterized by the quantity s. As the data in Table 11.2 clearly show, for the majority of cubic crystals this dimensionless quantity is significantly different from unity.

11.4.3 Elastic Waves in Cubic Crystals

Once the elastic properties are known, the propagation velocity of elastic waves in the crystal can be determined. We shall demonstrate this for waves propagating in directions of high symmetry in cubic crystals.

Starting with the equation

$$\rho\,\ddot{u}_\alpha = \sum_\beta \frac{\partial}{\partial r_\beta}\frac{\partial U_{\text{harm}}}{\partial \varepsilon_{\alpha\beta}} , \tag{11.4.44}$$

which is the consequence of (11.4.13) and (11.4.15), and using (11.4.35), the expression for elastic energy in cubic crystals, the equation that governs the propagation of elastic waves is written as

$$\begin{aligned}\rho\frac{\partial^2 u_x}{\partial t^2} = {} & c_{11}\frac{\partial^2 u_x}{\partial x^2} + c_{44}\left(\frac{\partial^2 u_x}{\partial y^2} + \frac{\partial^2 u_x}{\partial z^2}\right) \\ & + (c_{12} + c_{44})\left(\frac{\partial^2 u_y}{\partial x \partial y} + \frac{\partial^2 u_z}{\partial x \partial z}\right) ;\end{aligned} \tag{11.4.45}$$

similar equations apply to the y- and z-components of the displacement. Owing to the anisotropy of crystals, the propagation velocity of elastic waves depends on the crystallographic direction. For elastic waves propagating along the direction [100] strain is a function of the x-coordinate alone, consequently

$$\rho\frac{\partial^2 u_x}{\partial t^2} = c_{11}\frac{\partial^2 u_x}{\partial x^2}, \quad \rho\frac{\partial^2 u_y}{\partial t^2} = c_{44}\frac{\partial^2 u_y}{\partial x^2}, \quad \rho\frac{\partial^2 u_z}{\partial t^2} = c_{44}\frac{\partial^2 u_z}{\partial x^2} . \tag{11.4.46}$$

The propagation velocities for the longitudinal and two transverse waves of wave vector $\boldsymbol{q} = (q, 0, 0)$ are immediately read off:

$$c_{\mathrm{L}} = \left(\frac{c_{11}}{\rho}\right)^{1/2}, \qquad c_{\mathrm{T}} = \left(\frac{c_{44}}{\rho}\right)^{1/2}. \tag{11.4.47}$$

Note that this is the same result as for isotropic samples, provided Lamé constants are expressed in terms of the Voigt elastic constants c_{ij}.

For longitudinal waves of wave vector $\boldsymbol{q} = (q, q, q)/\sqrt{3}$ and propagation direction [111] the displacement is

$$\boldsymbol{u}_{\mathrm{L}}(\boldsymbol{r}) = \frac{1}{\sqrt{3}} \begin{pmatrix} u \\ u \\ u \end{pmatrix} \mathrm{e}^{\mathrm{i}(\boldsymbol{q}\cdot\boldsymbol{r}-\omega t)}, \tag{11.4.48}$$

while for the two transverse vibrations

$$\begin{aligned} \boldsymbol{u}_{\mathrm{T}_1}(\boldsymbol{r}) &= \frac{1}{\sqrt{2}} \begin{pmatrix} u \\ -u \\ 0 \end{pmatrix} \mathrm{e}^{\mathrm{i}(\boldsymbol{q}\cdot\boldsymbol{r}-\omega t)}, \\ \boldsymbol{u}_{\mathrm{T}_2}(\boldsymbol{r}) &= \frac{1}{\sqrt{6}} \begin{pmatrix} u \\ u \\ -2u \end{pmatrix} \mathrm{e}^{\mathrm{i}(\boldsymbol{q}\cdot\boldsymbol{r}-\omega t)}. \end{aligned} \tag{11.4.49}$$

Substituting these into the equations of motion, the same velocity is found for the two transverse branches, and a different one for the longitudinal:

$$c_{\mathrm{L}} = \left(\frac{c_{11} + 2c_{12} + 4c_{44}}{3\rho}\right)^{1/2}, \qquad c_{\mathrm{T}} = \left(\frac{c_{11} - c_{12} + c_{44}}{3\rho}\right)^{1/2}. \tag{11.4.50}$$

Along the propagation direction [110] the longitudinal wave of wave vector $\boldsymbol{q} = (q, q, 0)/\sqrt{2}$ is written as

$$\boldsymbol{u}_{\mathrm{L}}(\boldsymbol{r}) = \frac{1}{\sqrt{2}} \begin{pmatrix} u \\ u \\ 0 \end{pmatrix} \mathrm{e}^{\mathrm{i}(\boldsymbol{q}\cdot\boldsymbol{r}-\omega t)}, \tag{11.4.51}$$

while its transverse counterparts as

$$\boldsymbol{u}_{\mathrm{T}_1}(\boldsymbol{r}) = \frac{1}{\sqrt{2}} \begin{pmatrix} u \\ -u \\ 0 \end{pmatrix} \mathrm{e}^{\mathrm{i}(\boldsymbol{q}\cdot\boldsymbol{r}-\omega t)}, \qquad \boldsymbol{u}_{\mathrm{T}_2}(\boldsymbol{r}) = \begin{pmatrix} 0 \\ 0 \\ u \end{pmatrix} \mathrm{e}^{\mathrm{i}(\boldsymbol{q}\cdot\boldsymbol{r}-\omega t)}. \tag{11.4.52}$$

The propagation velocities are

$$c_{\mathrm{L}} = \left(\frac{c_{11} + c_{12} + 2c_{44}}{2\rho}\right)^{1/2}, \quad c_{\mathrm{T}_1} = \left(\frac{c_{11} - c_{12}}{2\rho}\right)^{1/2}, \quad c_{\mathrm{T}_2} = \left(\frac{c_{44}}{\rho}\right)^{1/2}. \tag{11.4.53}$$

Along this propagation direction the velocity is different for each branch.

The observation made for acoustic vibrations in simple cubic crystals applies to elastic waves, too: vibrations that are longitudinal and transverse with respect to the propagation direction exist only in some directions of sufficiently high symmetry. In the general case $\boldsymbol{u}(\boldsymbol{r})$ is neither parallel nor perpendicular to $\boldsymbol{q}$.

11.4.4 Optical Vibrations in Ionic Crystals

In contrast to acoustic vibrations, where atoms of the primitive cell oscillate in phase and with equal amplitudes in the long-wavelength limit, in optical vibrations atoms of the primitive cell oscillate completely out of phase and their center of mass remains stationary. Both longitudinal and transverse modes have finite frequencies, and they are different in general. In spite of this expectation, the longitudinal and transverse optical vibrations in silicon, shown in Fig. 11.12, are observed to be of the same frequency for $\boldsymbol{q} = 0$. That this is not accidental is confirmed by Fig. 11.13(a), where the same is observed in the measured vibrational spectrum of diamond.

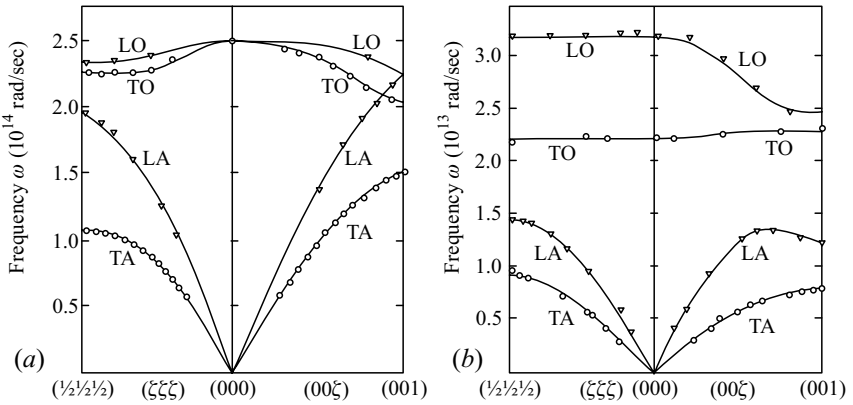


Fig. 11.13. (a) Dispersion relations for the lattice vibrations of diamond in two characteristic directions of the Brillouin zone [J. L. Warren et al., *Phys. Rev.* **158**, 805 (1967)]. (b) The same for NaI [based on A. D. B. Woods et al., *Phys. Rev.* **131**, 1025 (1963)]

Exploiting the symmetry properties of force constants it may be shown that in a cubic crystal longitudinal and transverse optical vibrations must indeed be of the same frequency at $\boldsymbol{q} = 0$. However, when the same measurements are performed in ionic crystals – results are shown in Fig. 11.13(b) and listed in Table 11.3 –, the two kinds of optical vibrations are found to be of unequal frequencies even in $\boldsymbol{q} = 0$.

Table 11.3. The frequency ν of optical vibrations (in units of 10^{12} Hz) at the center of the Brillouin zone for some covalently bonded elements and ionically bonded compounds that have two atoms per primitive cell

Element	$\nu_{\mathrm{LO}} = \nu_{\mathrm{TO}}$	Compound	ν_{LO}	ν_{TO}
C (diamond)	39.9	NaCl	7.91	4.92
Si	15.6	CsCl	4.95	2.97
Ge	9.0	GaAs	8.75	8.06

This indicates that, unlike in covalently bonded materials, the vibrational spectrum in ionic crystals cannot be represented by the mass–spring model. The underlying physical reason for the different behavior of ionic crystals is that the longitudinal motion of the atoms – in contrast to their transverse motion – gives rise to a periodically varying charge density, and through it to a periodically varying polarization. This generates an additional restoring force for longitudinal motions, and so such vibrations will be of higher frequency than transverse ones.

Since in the long-wavelength limit the motion of atoms is practically identical in every primitive cell, it is sufficient to examine the equations of motion in a single primitive cell. However, due account must be taken of the fact that inside ionic crystals a screened local electric field $\boldsymbol{E}_{\mathrm{eff}}$ acts on the atoms. Denoting the displacement of the ion of charge e^* $(-e^*)$ and mass M_+ (M_-) from its equilibrium position by $\boldsymbol{u}_+$ $(\boldsymbol{u}_-)$, the equations of motion are

$$\begin{aligned} M_+ \frac{\mathrm{d}^2 \boldsymbol{u}_+}{\mathrm{d}t^2} &= -K\,(\boldsymbol{u}_+ - \boldsymbol{u}_-) + e^* \boldsymbol{E}_{\mathrm{eff}}\,, \\ M_- \frac{\mathrm{d}^2 \boldsymbol{u}_-}{\mathrm{d}t^2} &= -K\,(\boldsymbol{u}_- - \boldsymbol{u}_+) - e^* \boldsymbol{E}_{\mathrm{eff}}\,. \end{aligned} \tag{11.4.54}$$

Combining these into a single equation for the relative displacement $\boldsymbol{u} = \boldsymbol{u}_+ - \boldsymbol{u}_-$:

$$\overline{M} \frac{\mathrm{d}^2 \boldsymbol{u}}{\mathrm{d}t^2} = -K\boldsymbol{u} + e^* \boldsymbol{E}_{\mathrm{eff}}\,, \tag{11.4.55}$$

where $\overline{M} = M_+ M_- / (M_+ + M_-)$ is the reduced mass.

In simple cubic crystals the local field and the macroscopic field are related by the Lorentz formula[6]

$$\boldsymbol{E}_{\mathrm{eff}} = \boldsymbol{E} + \frac{1}{3\epsilon_0} \boldsymbol{P}\,. \tag{11.4.56}$$

The polarization $\boldsymbol{P}$ comes from two sources. Denoting the polarizability of the medium due to the displacement of electrons by α, the local field induces a polarization

[6] H. A. Lorentz, 1880.

$$\boldsymbol{P}_{\mathrm{el}} = \frac{N}{V}\alpha \boldsymbol{E}_{\mathrm{eff}} \tag{11.4.57}$$

in a sample of volume V made up of N primitive cells. In addition to electronic polarization, an additional contribution arises from the displacement of ions, and so

$$\boldsymbol{P} = \frac{N}{V}(e^*\boldsymbol{u} + \alpha \boldsymbol{E}_{\mathrm{eff}})\,. \tag{11.4.58}$$

Substituting (11.4.56) into this equation, the polarization is

$$\boldsymbol{P} = \frac{N}{V}\frac{e^*\boldsymbol{u} + \alpha \boldsymbol{E}}{1 - (N\alpha/3V\epsilon_0)}\,. \tag{11.4.59}$$

In the high-frequency limit ions are unable to follow the rapid variations of the field, therefore they remain practically stationary, so

$$\boldsymbol{P} = \frac{N}{V}\frac{\alpha \boldsymbol{E}}{1 - (N\alpha/3V\epsilon_0)}\,. \tag{11.4.60}$$

Comparison with the general relationship

$$\boldsymbol{P}(\omega) = \epsilon(\omega)\boldsymbol{E} - \epsilon_0 \boldsymbol{E} = [\epsilon_{\mathrm{r}}(\omega) - 1]\epsilon_0 \boldsymbol{E} \tag{11.4.61}$$

between the polarization and the dielectric constant gives

$$\frac{N\alpha/V}{1 - (N\alpha/3V\epsilon_0)} = [\epsilon_{\mathrm{r}}(\infty) - 1]\epsilon_0\,, \tag{11.4.62}$$

which leads to a relation between the polarizability α and $\epsilon_{\mathrm{r}}(\infty)$ that is very similar to the *Clausius–Mossotti relation*:[7]

$$\frac{N}{3V}\alpha = \frac{\epsilon_{\mathrm{r}}(\infty) - 1}{\epsilon_{\mathrm{r}}(\infty) + 2}\,\epsilon_0\,. \tag{11.4.63}$$

Substituting this into (11.4.59), the expression for polarizability,

$$\boldsymbol{P} = \frac{N}{3V}e^*[\epsilon_{\mathrm{r}}(\infty) + 2]\boldsymbol{u} + [\epsilon_{\mathrm{r}}(\infty) - 1]\epsilon_0 \boldsymbol{E}\,. \tag{11.4.64}$$

Using this term in the Lorentz formula for the local field, and introducing the notation

$$\omega_0^2 = \frac{1}{\overline{M}}\left[K - \frac{N}{9V\epsilon_0}e^{*2}\big[\epsilon_{\mathrm{r}}(\infty) + 2\big]\right]\,, \tag{11.4.65}$$

the equation of motion (11.4.55) takes the form

$$\overline{M}\frac{\mathrm{d}^2\boldsymbol{u}}{\mathrm{d}t^2} = -\overline{M}\omega_0^2\boldsymbol{u} + \frac{1}{3}e^*\big[\epsilon_{\mathrm{r}}(\infty) + 2\big]\boldsymbol{E}\,. \tag{11.4.66}$$

[7] O.-F. Mossotti, 1850, and R. J. E. Clausius, 1879. When written in terms of the index of refraction, $n = \sqrt{\epsilon}$, it is known as the Lorentz–Lorenz equation (H. A. Lorentz and L. V. Lorenz, 1880).

In the static case the left-hand side of the equation vanishes. Substituting the relation between the static displacement and the electric field into the polarization formula (11.4.64), and making use of (11.4.61) at $\omega = 0$,

$$\epsilon_{\mathrm{r}}(0) - \epsilon_{\mathrm{r}}(\infty) = \frac{N}{9V\overline{M}\omega_0^2\epsilon_0} e^{*2} \left[\epsilon_{\mathrm{r}}(\infty) + 2\right]^2, \tag{11.4.67}$$

and therefore

$$\overline{M}\frac{\mathrm{d}^2\boldsymbol{u}}{\mathrm{d}t^2} = -\overline{M}\omega_0^2\boldsymbol{u} + \left(\frac{V\overline{M}\omega_0^2\epsilon_0}{N}\right)^{1/2} \left[\epsilon_{\mathrm{r}}(0) - \epsilon_{\mathrm{r}}(\infty)\right]^{1/2} \boldsymbol{E}\,. \tag{11.4.68}$$

Changing the variable $\boldsymbol{u}$ to

$$\boldsymbol{s} = \sqrt{\frac{N\overline{M}}{V}}\boldsymbol{u}\,, \tag{11.4.69}$$

the only remaining parameters in the equations of motion are the frequency ω_0 and the dielectric constant:

$$\frac{\mathrm{d}^2\boldsymbol{s}}{\mathrm{d}t^2} = -\omega_0^2\boldsymbol{s} + \omega_0\sqrt{\epsilon_0(\epsilon_{\mathrm{r}}(0) - \epsilon_{\mathrm{r}}(\infty))}\boldsymbol{E}\,. \tag{11.4.70}$$

In terms of the variable $\boldsymbol{s}$ the polarization reads

$$\boldsymbol{P} = \omega_0\sqrt{\epsilon_0(\epsilon_{\mathrm{r}}(0) - \epsilon_{\mathrm{r}}(\infty))}\boldsymbol{s} + \left[\epsilon_{\mathrm{r}}(\infty) - 1\right]\epsilon_0\boldsymbol{E}\,. \tag{11.4.71}$$

The quantity $\boldsymbol{s}$ is proportional to the displacement of ions. To decompose it into components parallel ($\boldsymbol{s}_{\mathrm{LO}}$) and perpendicular ($\boldsymbol{s}_{\mathrm{TO}}$) to the wave vector $\boldsymbol{q}$ – which specifies the propagation direction –, the conditions

$$\boldsymbol{q}\cdot\boldsymbol{s}_{\mathrm{TO}} = 0\,, \qquad \boldsymbol{q}\times\boldsymbol{s}_{\mathrm{LO}} = 0 \tag{11.4.72}$$

are exploited. Assuming wavelike spatial propagation with time-dependent coefficients for the longitudinal and transverse components alike,

$$\boldsymbol{s}_{\mathrm{LO}} = Q_{\mathrm{LO}}(t)\boldsymbol{e}_{\mathrm{LO}}\mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{R}_m}, \qquad \boldsymbol{s}_{\mathrm{TO}} = Q_{\mathrm{TO}}(t)\boldsymbol{e}_{\mathrm{TO}}\mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{R}_m}. \tag{11.4.73}$$

To solve the equation of motion, it should be noted that inside ionic crystals the macroscopic field satisfies the Maxwell equations

$$\mathrm{curl}\,\boldsymbol{E} = 0\,, \qquad \mathrm{div}\,\boldsymbol{D} = \mathrm{div}(\epsilon_0\boldsymbol{E} + \boldsymbol{P}) = 0\,. \tag{11.4.74}$$

Taking the curl of (11.4.70) gives

$$\frac{\mathrm{d}^2 Q_{\mathrm{TO}}}{\mathrm{d}t^2} = -\omega_0^2\, Q_{\mathrm{TO}} \tag{11.4.75}$$

for transverse vibrations; their frequency ω_{TO} is seen to be equal to ω_0.

To write up the equation for the longitudinal component, an equation derived from $\operatorname{div} \boldsymbol{D} = 0$ and (11.4.71) is used:

$$\operatorname{div} \boldsymbol{E} + \frac{\omega_0}{\epsilon_\mathrm{r}(\infty)} \sqrt{\frac{\epsilon_\mathrm{r}(0) - \epsilon_\mathrm{r}(\infty)}{\epsilon_0}} \operatorname{div} \boldsymbol{s} = 0 \,. \tag{11.4.76}$$

The normal coordinates of longitudinal optical vibrations must therefore satisfy

$$\frac{\mathrm{d}^2 Q_\mathrm{LO}}{\mathrm{d}t^2} = -\omega_0^2 \, \frac{\epsilon_\mathrm{r}(0)}{\epsilon_\mathrm{r}(\infty)} \, Q_\mathrm{LO} \,, \tag{11.4.77}$$

while the frequencies of LO vibrations are given by

$$\boxed{\omega_\mathrm{LO} = \left(\frac{\epsilon_\mathrm{r}(0)}{\epsilon_\mathrm{r}(\infty)} \right)^{1/2} \omega_\mathrm{TO} \,.} \tag{11.4.78}$$

This is the *Lyddane–Sachs–Teller relation.*[8] As we have seen, the difference between longitudinal and transverse optical frequencies is easily observed in ionic crystals, and $\omega_\mathrm{LO} > \omega_\mathrm{TO}$ is generally valid. The ratios calculated from relative dielectric constants measured at low and high frequencies are usually in good agreement with the ratios of the frequencies measured directly.

11.5 Localized Lattice Vibrations

It was mentioned in Chapter 9 in connection with the structure of real crystals that even in samples of the highest purity there are always foreign atoms, impurities. These obviously change the lattice dynamics, since both their mass and interactions with the neighbors are different from those of other atoms. Lattice vibrations are distorted around the impurity, hence their spectrum is also modified. To understand how an impurity may influence the spectrum of lattice vibrations, we shall discuss the simple case when a single atom of an ideal crystal is replaced by an impurity.

The mathematical description becomes rather complicated as the impurity breaks the translational symmetry of the crystal. Consequently, the states of the system cannot be characterized by the wave vector, as it is no longer a good quantum number. Nevertheless the frequency distribution of vibrational states remains a meaningful quantity.

11.5.1 Vibrations in a Chain with an Impurity

Once again, we shall use the example of a one-dimensional chain to demonstrate the method to account for the effects of an impurity. Consider a chain

[8] R. H. Lyddane, R. G. Sachs, and E. P. Teller, 1941.

made up of atoms of mass M, with the atom at site $n = 0$ replaced by one of different mass (M_0). For simplicity we shall assume that all force constants remain the same around the impurity. Following the steps of the discussion of monatomic linear chains, classical equations of motion are derived for the atomic displacements u_n. If only nearest neighbors interact, (11.2.3)

$$M\ddot{u}_n = -K\left[2u_n - u_{n-1} - u_{n+1}\right], \tag{11.5.1}$$

continues to be valid everywhere except for $n = 0$, where

$$M_0\ddot{u}_0 = -K\left[2u_0 - u_1 - u_{-1}\right] \tag{11.5.2}$$

because of the different mass of the atom. The two equations can be written jointly as

$$M\ddot{u}_n + K\left[2u_n - u_{n-1} - u_{n+1}\right] = \delta_{n,0}\,(M - M_0)\ddot{u}_0\,. \tag{11.5.3}$$

In spite of the impurity, vibrations are expected to propagate in the lattice; $u_n(t)$ is therefore Fourier transformed with respect to the time variable. The obtained Fourier coefficient, $u_n(\omega)$ then satisfies the equation

$$-\omega^2 M u_n + K\left[2u_n - u_{n+1} - u_{n-1}\right] = \delta_{n,0}\,\Delta M\omega^2 u_0\,, \tag{11.5.4}$$

where $\Delta M = M_0 - M$. Introducing the notation $\omega_0^2 = 4K/M$,

$$-\omega^2 u_n + \tfrac{1}{4}\omega_0^2\left[2u_n - u_{n+1} - u_{n-1}\right] = \delta_{n,0}\,(\Delta M/M)\omega^2 u_0\,. \tag{11.5.5}$$

Apart from the immediate vicinity of the impurity, atomic displacements of the form $u_n = A\lambda^n$ or $u_n = A\lambda^{-n}$ are sought. It is easily shown that λ is complex when $\omega < \omega_0$, and so the vibrations of the chain are wavelike. On the other hand, λ is real when $\omega > \omega_0$; vibrational amplitudes then decrease exponentially with the distance from the impurity, therefore only the impurity and its vicinity participate in the vibration.

Although the Hamiltonian of the system cannot be written as a sum of independent normal modes, $u_n(\omega)$ may be Fourier transformed with respect to the spatial variable as well, leading to the new variable

$$u(q,\omega) = \frac{1}{\sqrt{N}}\sum_n u_n(\omega)\mathrm{e}^{-\mathrm{i}qna}\,. \tag{11.5.6}$$

By Fourier transforming the quantity $F_n = \delta_{n,0}\,(\Delta M/M)\omega^2 u_0(\omega)$ on the right-hand side of (11.5.5), the equation of motion becomes

$$-\omega^2 u(q,\omega) + \tfrac{1}{4}\omega_0^2\left[2 - \mathrm{e}^{-\mathrm{i}qa} - \mathrm{e}^{\mathrm{i}qa}\right]u(q,\omega) = F(q)\,, \tag{11.5.7}$$

which can be simplified to

$$-\omega^2 u(q,\omega) + \omega_0^2\sin^2\left(\tfrac{1}{2}qa\right)u(q,\omega) = F(q)\,, \tag{11.5.8}$$

where

$$F(q) = \frac{1}{\sqrt{N}} \sum_n F_n \mathrm{e}^{-\mathrm{i}qna} = \frac{1}{\sqrt{N}} \frac{\Delta M}{M} \omega^2 u_0(\omega) \,. \tag{11.5.9}$$

Neglecting the expression on the right-hand side, the solutions of the equation are the previously determined eigenfrequencies $\omega(q) = \omega_0 \,|\sin(qa/2)|$ of the vibrations in a regular lattice. In terms of these the equation governing the vibrations of the lattice with an impurity reads

$$\left[\omega^2(q) - \omega^2\right] u(q,\omega) = F(q) \,. \tag{11.5.10}$$

Assuming that, owing to the impurity, each vibrational frequency is slightly shifted compared to the pure crystal, the formal solution of the equation for the vibrational amplitudes is

$$u(q,\omega) = \frac{F(q)}{\omega^2(q) - \omega^2} \,. \tag{11.5.11}$$

Fourier transforming this expression back into real space, and making use of the formula derived for $F(q)$,

$$u_n(\omega) = \frac{\Delta M}{M} \frac{1}{N} \sum_q \frac{\omega^2 u_0(\omega) \mathrm{e}^{\mathrm{i}qna}}{\omega^2(q) - \omega^2} \,. \tag{11.5.12}$$

The solution is self-consistent if the left-hand side gives the same u_0 for the displacement of the impurity atom at $n = 0$ as what appears on the right-hand side, in which case

$$\frac{\Delta M}{M} \frac{\omega^2}{N} \sum_q \frac{1}{\omega^2(q) - \omega^2} = 1 \,. \tag{11.5.13}$$

The solutions to this equation give the eigenfrequencies of the allowed vibrations of the chain with the impurity. To illustrate the equations graphically, it is first rearranged as

$$f(\omega^2) \equiv \frac{\omega^2}{N} \sum_q \frac{1}{\omega^2(q) - \omega^2} = \left(\frac{\Delta M}{M}\right)^{-1} \,. \tag{11.5.14}$$

The function $f(\omega^2)$ of the left-hand side is plotted against ω^2 in Fig. 11.14. The vibrational frequencies in the chain with an impurity are the horizontal coordinates (ω) of the intersection points of these curves with the horizontal line at $M/\Delta M$.

Since the left-hand side is discontinuous at each value $\omega^2(q)$ that corresponds to an eigenfrequency of the ideal chain, between any two adjacent eigenfrequencies there will be an intersection point, i.e., an allowed vibrational frequency of the chain with an impurity. If the impurity atom is heavier than the other atoms in the chain ($\Delta M > 0$) then each vibrational frequency is shifted slightly downward compared to the frequencies of the ideal chain, but

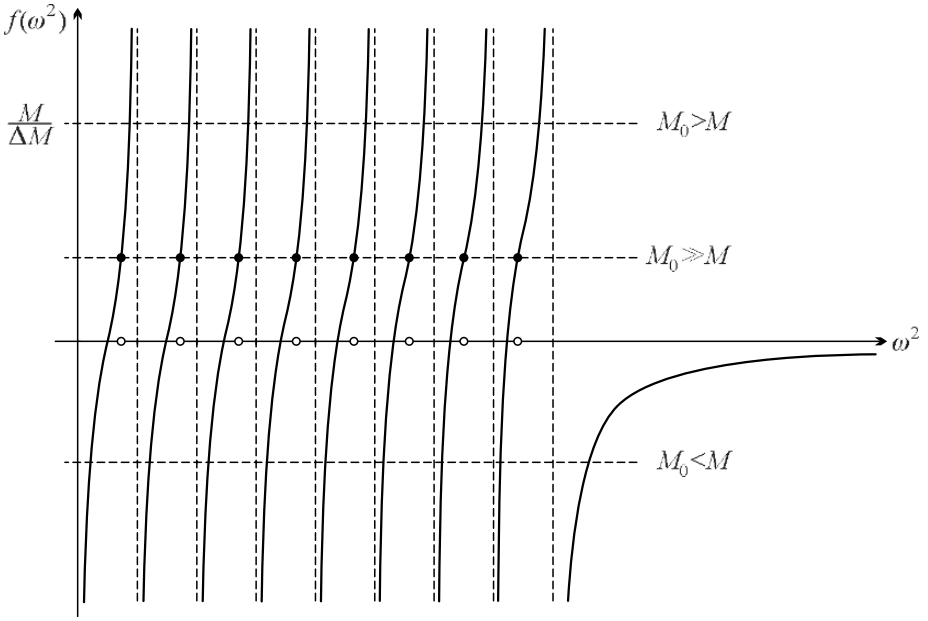


Fig. 11.14. Graphical determination of the eigenfrequencies of a chain with an impurity from the intersection points of $f(\omega^2)$ with the horizontal line at $M/\Delta M$

they cannot be lower than the next eigenfrequency. Although not visible in this equation, the original one nevertheless implies that $\omega = 0$ is also a solution, and it corresponds to a uniform translation for a chain with an impurity, too. Therefore vibrations fill – practically continuously – the same range as for a pure crystal. If the impurity atom is lighter than the others ($\Delta M < 0$) then the frequency of each vibrational mode is shifted slightly upward, but they cannot exceed that of the next eigenfrequency. The highest-frequency vibration is an exception, which appears as a separate mode above the continuum. Since $\omega > \omega_0$ for this vibration, the vibrational amplitude exponentially decreases with the distance from the impurity, as mentioned above. The vibration is localized, only the light impurity atom and its small vicinity participates in it with significant amplitudes. In diatomic chains such localized excitations may appear between the acoustic and optical branches or above the optical branch.

It is worth taking a closer look at the case when $M_0 > M$ and ΔM is comparable to the mass of the atoms. As we have seen, no localized vibration is possible in this case. Compared to the pure sample, vibrational frequencies are shifted downward, however, this shift is not uniform. Going back to the graphical solution of (11.5.14): if the right-hand side is of order unity, the frequencies determined by the intersection points are shifted only slightly with respect to the original ones at the low end of the spectrum, while at its high end

they are shifted significantly, almost to the frequency just below. Therefore somewhere around the middle of the spectrum frequencies are denser than elsewhere. Every atom of the crystal participates in these vibrations, however the displacement amplitude is much larger for the impurity than for distant atoms. It looks as if the impurity were resonating, therefore such states are called resonances.

11.5.2 Impurities in a Three-Dimensional Lattice

The previous method can now be generalized to three-dimensional lattices in which one atom is substituted by another atom of different mass, and thus force constants around the impurity are different from those in a regular crystal. The classical equation of motion (11.1.28) is replaced by a system of equations in which the mass $M_{m,\mu}$ of the μth atom in the mth primitive cell depends on the cell index m as well, and owing to the breakdown of translational symmetry force constants do not depend only on the separation of lattice points any more:

$$M_{m,\mu}\ddot{u}_\alpha(m,\mu,t) = -\sum_{n,\nu,\beta} \Phi^{\mu\nu}_{\alpha\beta}(m,n)u_\beta(n,\nu,t)\,. \tag{11.5.15}$$

Fourier transformation with respect to time gives

$$\omega^2 M_{m,\mu}u_\alpha(m,\mu,\omega) = \sum_{n,\nu,\beta} \Phi^{\mu\nu}_{\alpha\beta}(m,n)u_\beta(n,\nu,\omega)\,. \tag{11.5.16}$$

Rewriting this in the form

$$\sum_{n,\nu,\beta} \left[\omega^2 M_{m,\mu}\delta_{mn}\delta_{\mu\nu}\delta_{\alpha\beta} - \Phi^{\mu\nu}_{\alpha\beta}(m,n)\right] u_\beta(n,\nu,\omega) = 0\,, \tag{11.5.17}$$

the equation has nontrivial solutions if the matrix

$$L_{m,\mu,\alpha;n,\nu,\beta}(\omega^2) = \omega^2 M_{m,\mu}\delta_{mn}\delta_{\mu\nu}\delta_{\alpha\beta} - \Phi^{\mu\nu}_{\alpha\beta}(m,n) \tag{11.5.18}$$

made up of the coefficients of $u_\beta(n,\nu,\omega)$ has a vanishing determinant. In lattices without impurities, where translational invariance may be exploited, after a Fourier transformation with respect to the lattice points only a $3p\times 3p$ matrix needs to be diagonalized, since the equations for the Fourier coefficients of different $\boldsymbol{q}$s are not coupled. In contrast, when the crystal has an impurity, one has to deal with a $3pN\times 3pN$ matrix. Nevertheless similar statements can be made about the main features of the excitations as in the one-dimensional case.

For notational simplicity, we shall assume that the basis of the crystal is monatomic and that the impurity atom of mass M_0 sits at the lattice point $\boldsymbol{R}_m = 0$; moreover, we shall again neglect any modifications in the force constants. The equation of motion for the impurity atom is then

$$M_0\ddot{u}_\alpha(0) = -\sum_{n,\beta} \Phi_{\alpha\beta}(n)u_\beta(n)\,. \tag{11.5.19}$$

Using the notation $\Delta M = M_0 - M$, the equation of motion can be rewritten as

$$M\ddot{u}_\alpha(0) + \sum_{n,\beta} \Phi_{\alpha\beta}(n)u_\beta(n) + \Delta M\ddot{u}_\alpha(0) = 0\,. \tag{11.5.20}$$

Together with the equations of motion for other atoms, this can be written in the common form

$$M\ddot{u}_\alpha(m) + \sum_{n,\beta} \Phi_{\alpha\beta}(m-n)u_\beta(n) + \delta_{m0}\Delta M\ddot{u}_\alpha(0) = 0\,, \tag{11.5.21}$$

or, using the Fourier transforms with respect to time,

$$\sum_{n,\beta} \left[\omega^2 M\delta_{mn}\delta_{\alpha\beta} - \Phi_{\alpha\beta}(m-n)\right] u_\beta(n) + \delta_{m0}\omega^2\Delta M u_\alpha(0) = 0\,. \tag{11.5.22}$$

The expression for matrix L is then

$$L_{m,\alpha;n,\beta}(\omega^2) = \omega^2 M\delta_{mn}\delta_{\alpha\beta} - \Phi_{\alpha\beta}(m-n) + \delta_{\alpha\beta}\delta_{m0}\omega^2\Delta M\,. \tag{11.5.23}$$

The first two terms on the right-hand side are recognized as the matrix L^0 that governs the vibrations of pure crystals. Separating this leads to

$$L_{m,\alpha;n,\beta}(\omega^2) = L^0_{m,\alpha;n,\beta}(\omega^2) + \delta L_{m,\alpha;n,\beta}(\omega^2)\,, \tag{11.5.24}$$

where

$$\delta L_{m,\alpha;n,\beta}(\omega^2) = \delta_{m0}\delta_{\alpha\beta}\omega^2\Delta M\,. \tag{11.5.25}$$

Next the inverse of L^0 is introduced through the definition

$$\sum_{m',\alpha'} R_{m,\alpha;m',\alpha'}(\omega^2)L^0_{m',\alpha';n,\beta}(\omega^2) = \delta_{mn}\delta_{\alpha\beta}\,. \tag{11.5.26}$$

Since the eigenvectors of matrix L^0 are the polarization vectors and its eigenfrequencies are the vibrational frequencies of the pure crystal, it is easily seen that

$$R_{m,\alpha;n,\beta}(\omega^2) = \frac{1}{MN}\sum_{\boldsymbol{q}\lambda} \frac{\boldsymbol{e}^{(\lambda)^*}_\alpha(\boldsymbol{q})\boldsymbol{e}^{(\lambda)}_\beta(\boldsymbol{q})}{\omega^2 - \omega^2_\lambda(\boldsymbol{q})} \mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot(\boldsymbol{R}_m - \boldsymbol{R}_n)}\,. \tag{11.5.27}$$

The equation of motion (11.5.22) may be written in the concise form $(L^0 + \delta L)u = 0$. Multiplying it from the left by R,

$$(1 + R\cdot\delta L)u = 0 \tag{11.5.28}$$

is obtained. Using the above forms of R and δL, the equation governing the displacement of the impurity atom is

$$\left[1+\frac{\Delta M}{M}\frac{\omega^2}{N}\sum_{\boldsymbol{q}\lambda}\frac{\boldsymbol{e}_\alpha^{(\lambda)^*}(\boldsymbol{q})\boldsymbol{e}_\alpha^{(\lambda)}(\boldsymbol{q})}{\omega^2-\omega_\lambda^2(\boldsymbol{q})}\right]u_\alpha(0)=0\,. \tag{11.5.29}$$

The condition of self-consistency is now

$$\frac{\Delta M}{M}\frac{\omega^2}{N}\sum_{\boldsymbol{q}\lambda}\frac{\boldsymbol{e}^{(\lambda)^*}(\boldsymbol{q})\boldsymbol{e}^{(\lambda)}(\boldsymbol{q})}{\omega_\lambda^2(\boldsymbol{q})-\omega^2}=1\,. \tag{11.5.30}$$

Its structure is very similar to that of the condition obtained for a one-dimensional chain with an impurity. Using similar graphical methods for determining the solutions, vibrational frequencies are usually found to lie inside the original quasicontinuum, slightly shifted with respect to the unperturbed values. For heavy impurities the deformation of the spectrum is such that a resonance may show up inside it, while for light impurities a frequency that corresponds to a localized vibration may appear above the continuum. If the Debye spectrum (to be discussed in the next chapter) is used for characterizing the unperturbed lattice, such a localized vibration is observed to appear for any $\Delta M < 0$. Numerical calculations show that when a more realistic spectrum is chosen, localized vibrations show up only above a certain mass difference.

11.6 The Specific Heat of Classical Lattices

Having established that the thermal motion of ions about their equilibrium positions may be described in terms of normal modes, it is a relatively simple matter to determine the thermal properties of a classical crystal. According to statistical mechanics, the thermal energy of the crystal may be derived from the formula

$$E=\frac{\int \mathrm{d}\Gamma\,\mathcal{H}\mathrm{e}^{-\beta\mathcal{H}}}{\int \mathrm{d}\Gamma\,\mathrm{e}^{-\beta\mathcal{H}}}\,, \tag{11.6.1}$$

where integration is over the phase space spanned by the variables $\boldsymbol{u}(m,\mu)$ and $\boldsymbol{P}(m,\mu)$, that is

$$\mathrm{d}\Gamma=\prod_{m,\mu}\mathrm{d}\boldsymbol{u}(m,\mu)\,\mathrm{d}\boldsymbol{P}(m,\mu)\,. \tag{11.6.2}$$

Calculations are performed more easily when displacements and momenta are expressed in terms of normal coordinates. Using the form (11.3.30) for the Hamiltonian, the volume element in the phase space (Q,P) is

$$\mathrm{d}\Gamma=\prod_{\boldsymbol{q},\lambda}\mathrm{d}|Q_\lambda(\boldsymbol{q})|\,\mathrm{d}|P_\lambda(\boldsymbol{q})|\,. \tag{11.6.3}$$

The total energy is the sum over all independent modes,

$$E = \sum_{\boldsymbol{q},\lambda} \frac{\int \mathrm{d}\Gamma_\lambda(\boldsymbol{q})\, \mathcal{H}_\lambda(\boldsymbol{q}) \mathrm{e}^{-\beta \mathcal{H}_\lambda(\boldsymbol{q})}}{\int \mathrm{d}\Gamma_\lambda(\boldsymbol{q})\, \mathrm{e}^{-\beta \mathcal{H}_\lambda(\boldsymbol{q})}}\,, \tag{11.6.4}$$

where, in accordance with (11.3.30),

$$\mathcal{H}_\lambda(\boldsymbol{q}) = \tfrac{1}{2}\left\{|P_\lambda(\boldsymbol{q})|^2 + \omega_\lambda^2(\boldsymbol{q})|Q_\lambda(\boldsymbol{q})|^2\right\}. \tag{11.6.5}$$

Owing to their quadratic form, the contributions of the kinetic and potential energies of a vibrational mode are both $k_\mathrm{B}T/2$, as established by the equipartition theorem in statistical physics. Since each of the $3Np$ possible modes contribute $k_\mathrm{B}T$,

$$E = 3Np\, k_\mathrm{B} T\,. \tag{11.6.6}$$

The contribution of ionic vibrations to the specific heat is then

$$C_V = \frac{\partial E}{\partial T} = 3Np\, k_\mathrm{B}\,. \tag{11.6.7}$$

For the molar heat (specific heat per mole) of monatomic solids this gives the classical result expressed by the *Dulong–Petit law*:[9]

$$C_V = 3pR = 24.943\,\mathrm{J\,mol^{-1}\,K^{-1}} = 5.958\,\mathrm{cal\,mol^{-1}\,K^{-1}}. \tag{11.6.8}$$

Around and above room temperature, the measured value of the specific heat is close to the Dulong–Petit value in most cases. By the end of the 19th century it had become clear that at lower temperatures the specific heat drops. Today it is also known that it vanishes at the absolute zero of temperature, as shown in Fig. 11.15.

This departure from the theoretical value casts doubts on the applicability of the used approximations. Two important approximations were made: the

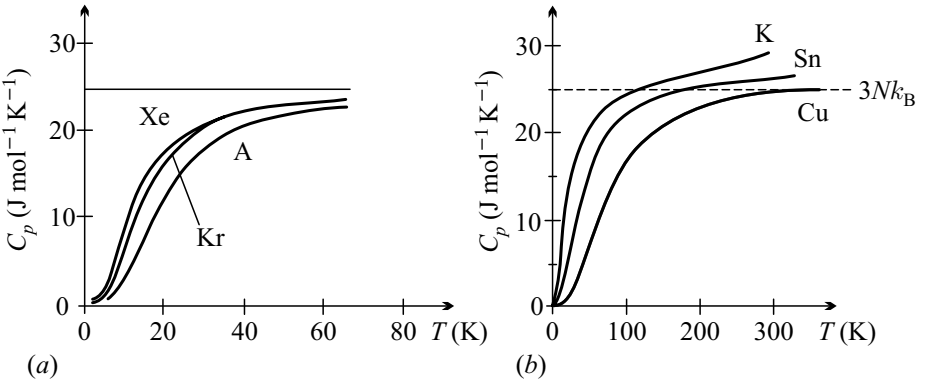


Fig. 11.15. Temperature dependence of the specific heat for (*a*) noble-gas crystals; (*b*) some metals

[9] P. L. Dulong and A. T. Petit, 1819.

harmonic expansion of the potential and the classical treatment of the vibrations. In the next chapter we shall first examine the role of quantum effects, and then analyze the role of the terms beyond the harmonic approximation.

Further Reading

1. M. Born and K. Huang, *Dynamical Theory of Crystal Lattices*, Oxford Classic Texts in the Physical Sciences, Oxford University Press, Oxford (1998).
2. P. Brüesch, *Phonons: Theory and Experiments I, Lattice Dynamics and Models of Interatomic Forces*, Springer-Verlag, Berlin (1982).
3. A. A. Maradudin, E. W. Montroll, G. H. Weiss, I. P. Ipatova, *Theory of Lattice Dynamics in the Harmonic Approximation*, Second Edition, in: Solid State Physics, Supplement 3, Academic Press, New York (1971).
4. J. A. Reissland, *The Physics of Phonons*, John Wiley & Sons Ltd., London (1973).

12

The Quantum Theory of Lattice Vibrations

In the previous chapter lattice vibrations were analyzed using the classical equations of motion. It was established that a much higher specific heat is predicted by classical statistical physics for the thermal excitation of such modes than what is observed experimentally at low temperatures. To obtain a better agreement a quantum mechanical treatment is required. In addition to providing a more precise description for the thermodynamic behavior of solids, this will also help to understand the influence of lattice vibrations on transport and optical properties (which will be discussed in Volume 2).

12.1 Quantization of Lattice Vibrations

Perhaps the most important result of the previous chapter is that the classical oscillatory motion of atoms in a lattice can be described in terms of harmonic oscillators. Starting with this, it is a fairly simple matter to derive and illustrate the quantized energy spectrum of a vibrating lattice, as the determination of the wavefunction and energy eigenvalues of a harmonic oscillator ranks among the easiest problems in quantum mechanics. Before turning to the general solution of the quantum mechanical problem, we shall present some simple models.

12.1.1 The Einstein Model

Shortly after PLANCK's proposal of the quantum hypothesis,[1] in 1907 EINSTEIN extended the hypothesis to lattice vibrations, postulating that their energy is also quantized, and for a lattice vibration of frequency ν it can

[1] M. PLANCK, 1900. MAX KARL ERNST LUDWIG PLANCK (1858–1947) was awarded the Nobel Prize in 1918 "in recognition of the services he rendered to the advancement of Physics by his discovery of energy quanta".

take only integral multiples of $h\nu$: $\varepsilon_n = nh\nu$. The mean thermal energy of a vibrational mode is then

$$\langle\varepsilon\rangle = \frac{\sum_{n=0}^{\infty} \varepsilon_n \mathrm{e}^{-\varepsilon_n/k_\mathrm{B}T}}{\sum_{n=0}^{\infty} \mathrm{e}^{-\varepsilon_n/k_\mathrm{B}T}} = \frac{\sum_{n=0}^{\infty} nh\nu \mathrm{e}^{-nh\nu/k_\mathrm{B}T}}{\sum_{n=0}^{\infty} \mathrm{e}^{-nh\nu/k_\mathrm{B}T}} \,. \tag{12.1.1}$$

In terms of the partition function

$$Z = \sum_{n=0}^{\infty} \mathrm{e}^{-\varepsilon_n/k_\mathrm{B}T} \tag{12.1.2}$$

the mean thermal energy may be written as

$$\langle\varepsilon\rangle = -\frac{\mathrm{d}}{\mathrm{d}\beta} \ln Z \,, \tag{12.1.3}$$

where $\beta = 1/k_\mathrm{B}T$. For harmonic oscillators the partition function is the sum of an infinite geometrical series:

$$Z = \sum_{n=0}^{\infty} \mathrm{e}^{-nh\nu/k_\mathrm{B}T} = \frac{1}{1 - \mathrm{e}^{-h\nu/k_\mathrm{B}T}} \,, \tag{12.1.4}$$

and so

$$\langle\varepsilon\rangle = \frac{h\nu}{\mathrm{e}^{h\nu/k_\mathrm{B}T} - 1} \,. \tag{12.1.5}$$

For simplicity Einstein also assumed that independently of each other, atoms oscillate with the same frequency ν_E; in other words frequency is independent of the wave vector of lattice vibrations. The model based on the assumption that the frequency is the same for each vibrational mode is called the *Einstein model*. Instead of the frequency ν_E we shall use the angular frequency ω_E. Assuming that there are Np atoms in the crystal, and each of them can oscillate in any of the three spatial directions, the internal energy of these vibrations is

$$E = 3Np\,\langle\varepsilon\rangle = 3Np \frac{\hbar\omega_\mathrm{E}}{\mathrm{e}^{\hbar\omega_\mathrm{E}/k_\mathrm{B}T} - 1} \,. \tag{12.1.6}$$

From this expression the specific heat due to lattice vibrations is found to be

$$C_V = \frac{\partial E}{\partial T} = 3Np\,k_\mathrm{B} \left(\frac{\hbar\omega_\mathrm{E}}{k_\mathrm{B}T}\right)^2 \frac{\mathrm{e}^{\hbar\omega_\mathrm{E}/k_\mathrm{B}T}}{(\mathrm{e}^{\hbar\omega_\mathrm{E}/k_\mathrm{B}T} - 1)^2} \,. \tag{12.1.7}$$

On the right-hand side

$$F(x) = \frac{x^2 \mathrm{e}^x}{(\mathrm{e}^x - 1)^2} \tag{12.1.8}$$

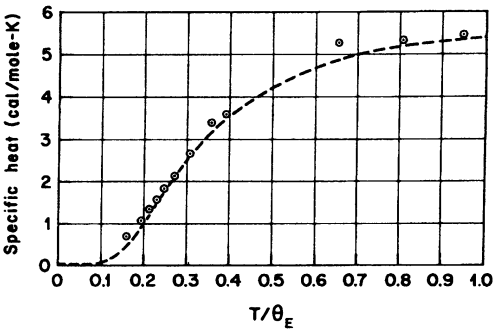


Fig. 12.1. The temperature dependence of specific heat in the Einstein model, fitted to experimental data for diamond [based on A. Einstein, *Ann. d. Phys.* **22**, 180 (1907)]

is the Einstein function with the argument $x = \hbar\omega_{\mathrm{E}}/k_{\mathrm{B}}T$. Figure 12.1 shows the temperature dependence of the specific heat, compared to the experimental data that were known at the beginning of the 20th century.

At sufficiently high temperatures, where $k_{\mathrm{B}}T \gg \hbar\omega_{\mathrm{E}}$, the first term of the expansion of the Einstein function for small values of x,

$$F(x) = \frac{x^2(1 + x + \dots)}{(x + x^2/2 + \dots)^2} = 1 + \mathcal{O}(x^2) \tag{12.1.9}$$

leads to the Dulong–Petit value of specific heat. At lower temperatures finite-energy oscillations are more difficult to excite thermally, therefore the specific heat decreases. In the $T \to 0$ limit asymptotically

$$F(x) = x^2 \mathrm{e}^{-x}, \qquad \text{if} \quad x \gg 1\,, \tag{12.1.10}$$

that is, the specific heat vanishes exponentially. Measurements performed since that time in the low-temperature region have shown a slower, power-law decay. Obviously, the excessively rapid decay of the specific heat in the Einstein model is due to the absence of long-wavelength acoustic vibrations that correspond to low-frequency – and therefore low-energy – collective modes. Such modes can be excited thermally more easily than Einstein oscillators.

12.1.2 The Debye Model

To correct the error of the Einstein model mentioned above, in 1912 P. Debye proposed that instead of individual atomic oscillations, elastic waves propagating in the solid with linear dispersion relation should be considered, and quantized according to Planck's hypothesis. Thus Debye assumed that the energy of compressional (longitudinal) and torsional (transverse) waves of angular frequency $\omega_{\mathrm{L}}(\boldsymbol{q})$ ($\omega_{\mathrm{T}}(\boldsymbol{q})$) can only change in steps of $\hbar\omega_{\mathrm{L}}(\boldsymbol{q})$ ($\hbar\omega_{\mathrm{T}}(\boldsymbol{q})$).

If it is possible to excite an arbitrary number of such vibrations, the thermal energy associated with the vibration of polarization λ and wave vector $\boldsymbol{q}$ can be determined along the same lines as in the Einstein model. Making use of the result

$$\langle \varepsilon_\lambda(\boldsymbol{q}) \rangle = \frac{\hbar\omega_\lambda(\boldsymbol{q})}{\mathrm{e}^{\hbar\omega_\lambda(\boldsymbol{q})/k_\mathrm{B}T} - 1}\,, \tag{12.1.11}$$

the total thermal energy of the crystal is

$$E = \sum_{\boldsymbol{q},\lambda} \frac{\hbar\omega_\lambda(\boldsymbol{q})}{\mathrm{e}^{\hbar\omega_\lambda(\boldsymbol{q})/k_\mathrm{B}T} - 1}\,. \tag{12.1.12}$$

By extending the classical results for elastic continua Debye assumed that the strict proportionality between $\omega_\lambda(\boldsymbol{q})$ and the wave number holds not only in the long-wavelength region but all the way down to wavelengths on the order of atomic dimensions. For simplicity, he also assumed that longitudinal and transverse vibrations propagate with the same velocity, and that the allowed values of the vectors $\boldsymbol{q}$ fill a sphere in reciprocal space rather than the Brillouin zone determined by the specific crystal structure. As there are one longitudinal and two transverse modes for any $\boldsymbol{q}$, the radius q_D of this sphere has to be chosen such that three times the number of allowed $\boldsymbol{q}$ vectors in its interior should be equal to the number of normal modes, $3pN$. Since a reciprocal-space volume $(2\pi)^3/V$ is associated with each vector $\boldsymbol{q}$,

$$\frac{4\pi q_\mathrm{D}^3}{3}\frac{V}{(2\pi)^3} = pN\,. \tag{12.1.13}$$

In crystals with a polyatomic basis excitations in the optical branches are taken into account by larger a q_D than in the monatomic case.

These are the underlying assumptions of the *Debye model*. Before turning to the detailed discussion of its properties, we shall present a more precise formulation of quantization.

12.1.3 Quantization of the Hamiltonian

In the Einstein model the solid was regarded as a collection of atoms that vibrate at the same frequency and are independent in a sense. In contrast, in the Debye model elastic waves propagating in a continuous elastic medium are considered, and the atomic structure of the solid is ignored. Looking back at the classical description of lattice vibrations presented in the previous chapter, one may say that the Einstein model describes optical modes, while the elastic waves in the Debye model correspond to the long-wavelength limit of acoustic waves.

Depending on the basis, optical and acoustic vibrations may be simultaneously present in real crystals; the dispersion relation is then different from the one used in the previous approximation. Through plausible assumptions

about the potential and the force constants the entire classical vibrational spectrum may be determined in principle, it is thus straightforward to suppose that Planck's quantum hypothesis should be applied to these frequencies $\omega_\lambda(\boldsymbol{q})$. The quantization condition can be formulated by stipulating that, for acoustic and optical vibrations alike, regardless of the dispersion relation, the energy of a vibration of frequency $\omega_\lambda(\boldsymbol{q})$ can change only in quanta of $\hbar\omega_\lambda(\boldsymbol{q})$, but the number of energy quanta can be arbitrarily large.

This also means that (12.1.11) and (12.1.12) are accepted for any dispersion relation. As we shall see, up to a temperature-independent constant term this yields the correct result for the total energy. Using (12.1.12), the thermal energy – and from it, the thermodynamic properties of the crystal – may then be determined, at least in principle, when the classical vibrational spectrum is known. To determine and physically interpret the additive constant, a more precise definition of quantization is required.

In the previous chapter we saw that using the canonically conjugate variables $\boldsymbol{u}(m,\mu)$ and $\boldsymbol{P}(m,\mu)$, the classical Hamiltonian (11.1.25) can be written as the sum of the kinetic energy (11.1.24) and the potential energy (11.1.21). Transition to quantum mechanics is then straightforward through canonical quantization. To obtain the quantum mechanical Hamiltonian operator from the classical Hamiltonian, the variables $\boldsymbol{u}$ and $\boldsymbol{P}$ in T_{kin} and U_{harm} are considered as operators, and the canonical commutation relation

$$[u_\alpha(m,\mu), P_\beta(n,\nu)] = \mathrm{i}\hbar\delta_{\alpha\beta}\delta_{mn}\delta_{\mu\nu} \tag{12.1.14}$$

is imposed. This is automatically satisfied when expression

$$\boldsymbol{P}(m,\mu) = \frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial \boldsymbol{u}(m,\mu)} \tag{12.1.15}$$

is used.

Repeating the steps that led – through the solution of the eigenvalue problem of the dynamical matrix – from the displacements $\boldsymbol{u}(m,\mu)$ to the normal coordinates, $Q_\lambda(\boldsymbol{q})$ also appears as an operator: with its canonically conjugate momentum $P_\lambda(\boldsymbol{q})$ it satisfies the commutation relations

$$\begin{aligned} [Q_\lambda(\boldsymbol{q}), P_{\lambda'}(\boldsymbol{q}')] &= \mathrm{i}\hbar\delta_{\lambda\lambda'}\delta_{\boldsymbol{q}\boldsymbol{q}'}\,, \\ [Q_\lambda(\boldsymbol{q}), Q_{\lambda'}(\boldsymbol{q}')] &= [P_\lambda(\boldsymbol{q}), P_{\lambda'}(\boldsymbol{q}')] = 0\,. \end{aligned} \tag{12.1.16}$$

The Hamiltonian then reads

$$\mathcal{H}_{\mathrm{harm}} = \tfrac{1}{2}\sum_{\boldsymbol{q},\lambda}\left\{P_\lambda^\dagger(\boldsymbol{q})P_\lambda(\boldsymbol{q}) + \omega_\lambda^2(\boldsymbol{q})Q_\lambda^\dagger(\boldsymbol{q})Q_\lambda(\boldsymbol{q})\right\}\,, \tag{12.1.17}$$

where $Q_\lambda^\dagger(\boldsymbol{q})$ and $P_\lambda^\dagger(\boldsymbol{q})$ are the Hermitian conjugates of $Q_\lambda(\boldsymbol{q})$ and $P_\lambda(\boldsymbol{q})$. Note that the same result is obtained when one starts with the classical Hamiltonian (11.3.30) expressed with the normal coordinates, and changes

to a quantum mechanical description by imposing the canonical commutation relations between the normal coordinates $Q_\lambda(\boldsymbol{q})$ and their canonically conjugate momenta $P_\lambda(\boldsymbol{q})$.

The quantum mechanical Hamiltonian is then the sum of the Hamiltonians of $3pN$ independent quantum oscillators. Nevertheless the vibrational frequencies can still be determined classically, as they are the eigenvalues of the dynamical matrix, the second derivative of the potential.

12.1.4 The Quantum Mechanics of Harmonic Oscillators

To gain insight into the properties of solids that are due to lattice vibrations, the quantum mechanics of harmonic oscillators has to be recalled. Before absorbing the mass into the normal coordinates, the Hamiltonian reads

$$\mathcal{H} = \frac{p^2}{2m} + \tfrac{1}{2}m\omega^2 x^2 = -\frac{\hbar^2}{2m}\left(\frac{\mathrm{d}}{\mathrm{d}x}\right)^2 + \tfrac{1}{2}m\omega^2 x^2. \tag{12.1.18}$$

Imposing the requirement of square integrability – whereby the wavefunction vanishes sufficiently rapidly at infinity –, the eigenvalues of this differential operator are expressed in terms of the Hermite polynomials (presented in Appendix C). The eigenfunctions normalized to unity are

$$\psi_n(x) = \left(\frac{m\omega}{\hbar\pi}\right)^{1/4} (2^n\, n!)^{-1/2} H_n\left(\sqrt{\frac{m\omega}{\hbar}}x\right) \mathrm{e}^{-m\omega x^2/2\hbar}, \tag{12.1.19}$$

and the energy eigenvalue of the state of quantum number n is

$$\varepsilon_n = \hbar\omega\left(n + \tfrac{1}{2}\right). \tag{12.1.20}$$

The wavefunction of the ground state is a Gaussian of finite width,

$$\psi_0(x) = \left(\frac{m\omega}{\hbar\pi}\right)^{1/4} \mathrm{e}^{-m\omega x^2/2\hbar}. \tag{12.1.21}$$

According to the uncertainty principle, the oscillator performs small-amplitude vibrations, so-called *zero-point vibrations* even in its ground state. This is indicated by the finite value of the mean square displacement,

$$\langle x^2\rangle = \left(\frac{m\omega}{\hbar\pi}\right)^{1/2} \int_{-\infty}^{\infty} x^2 \mathrm{e}^{-m\omega x^2/\hbar}\,\mathrm{d}x = \frac{\hbar}{2m\omega}. \tag{12.1.22}$$

This is why the ground-state energy is finite, $\varepsilon_0 = \frac{1}{2}\hbar\omega$; half of it comes from the kinetic energy and the other half from the potential energy of the oscillator. This finite ground-state energy is called the *zero-point energy.*

The solutions of the eigenvalue problem were determined in terms of the Hermitian operators x and p, which satisfy canonical commutation relations.

An easier alternative is offered by writing the Hamiltonian in terms of the non-Hermitian operator a and its adjoint defined as

$$a = \sqrt{\frac{m\omega}{2\hbar}}\left(x + \frac{\mathrm{i}}{m\omega}p\right), \qquad a^\dagger = \sqrt{\frac{m\omega}{2\hbar}}\left(x - \frac{\mathrm{i}}{m\omega}p\right). \tag{12.1.23}$$

Using the new operators, x and p are expressed as

$$x = \sqrt{\frac{\hbar}{2m\omega}}\left(a + a^\dagger\right), \qquad p = \mathrm{i}\sqrt{\frac{\hbar m\omega}{2}}\left(a^\dagger - a\right). \tag{12.1.24}$$

Making use of the canonical commutation relation $[x, p] = \mathrm{i}\hbar$, it is immediately established that

$$[a, a^\dagger] = \frac{\mathrm{i}}{\hbar}(px - xp) = 1, \tag{12.1.25}$$

and the Hamiltonian is

$$\mathcal{H} = \tfrac{1}{2}\hbar\omega\left(a^\dagger a + aa^\dagger\right) = \hbar\omega\left(a^\dagger a + \tfrac{1}{2}\right). \tag{12.1.26}$$

The energy eigenstates and eigenvalues are thus related to the eigenfunctions and eigenvalues of the operator $\hat{n} = a^\dagger a$. The commutation relation

$$[\hat{n}, a] = -a, \qquad [\hat{n}, a^\dagger] = a^\dagger \tag{12.1.27}$$

implies that if ψ_n is an eigenfunction of operator $\hat{n}$ with eigenvalue n, then the $a\psi_n$ and $a^\dagger\psi_n$ are also eigenfunctions, with eigenvalues $n - 1$ and $n + 1$:

$$\hat{n}\big(a\psi_n\big) = (n-1)a\psi_n, \qquad \hat{n}\big(a^\dagger\psi_n\big) = (n+1)a^\dagger\psi_n. \tag{12.1.28}$$

For this reason a and $a^\dagger$ are called *ladder operators* or *shift operators*. To obtain, as expected on physical grounds, a spectrum that is bounded from below, the eigenvalues of $\hat{n}$ have to be integers, and the lowest of them zero. The eigenvalues of $\hat{n}$ are then nonnegative integers,

$$\hat{n}\psi_n = n\psi_n \qquad n \in \mathbf{N}_0. \tag{12.1.29}$$

Energy eigenvalues are indeed given by (12.1.20). The restriction that the quantum number n can take only nonnegative integers justifies the picture that in state ψ_n there are n quanta of energy $\hbar\omega$ present, and $a^\dagger$ is their creation operator while a is their annihilation (destruction) operator.

Starting with the ground-state wavefunction ψ_0, every eigenstate can be constructed by means of the creation operator $a^\dagger$. The state of quantum number n that is normalized to unity is

$$\psi_n = \frac{1}{\sqrt{n!}}\left(a^\dagger\right)^n\psi_0. \tag{12.1.30}$$

Adding a further quantum to this state or taking away one from it,

$$a^\dagger\psi_n = \sqrt{n+1}\,\psi_{n+1}, \qquad a\,\psi_n = \sqrt{n}\,\psi_{n-1}. \tag{12.1.31}$$

12.1.5 Creation and Annihilation Operators of Vibrational Modes

By generalizing the results of the quantum mechanical treatment of the harmonic oscillator, we shall introduce the shift operators for the oscillator that corresponds to a lattice vibration of polarization λ and wave vector $\boldsymbol{q}$:

$$\begin{aligned} a_\lambda(\boldsymbol{q}) &= \sqrt{\frac{\omega_\lambda(\boldsymbol{q})}{2\hbar}}\left(Q_\lambda(\boldsymbol{q}) + \frac{\mathrm{i}}{\omega_\lambda(\boldsymbol{q})}P_\lambda^\dagger(\boldsymbol{q})\right) \\ &= \sqrt{\frac{\omega_\lambda(\boldsymbol{q})}{2\hbar}}\left(Q_\lambda(\boldsymbol{q}) + \frac{\mathrm{i}}{\omega_\lambda(\boldsymbol{q})}P_\lambda(-\boldsymbol{q})\right), \qquad (12.1.32\text{-a}) \\ a_\lambda^\dagger(\boldsymbol{q}) &= \sqrt{\frac{\omega_\lambda(\boldsymbol{q})}{2\hbar}}\left(Q_\lambda^\dagger(\boldsymbol{q}) - \frac{\mathrm{i}}{\omega_\lambda(\boldsymbol{q})}P_\lambda(\boldsymbol{q})\right) \\ &= \sqrt{\frac{\omega_\lambda(\boldsymbol{q})}{2\hbar}}\left(Q_\lambda(-\boldsymbol{q}) - \frac{\mathrm{i}}{\omega_\lambda(\boldsymbol{q})}P_\lambda(\boldsymbol{q})\right). \qquad (12.1.32\text{-b}) \end{aligned}$$

It follows immediately from the commutation relations between normal coordinates and their conjugate momenta that these operators satisfy the bosonic commutation relations

$$\begin{aligned} \left[a_\lambda(\boldsymbol{q}), a_{\lambda'}^\dagger(\boldsymbol{q}')\right] &= \delta_{\lambda,\lambda'}\,\delta_{\boldsymbol{q},\boldsymbol{q}'}\,, \\ \left[a_\lambda(\boldsymbol{q}), a_{\lambda'}(\boldsymbol{q}')\right] &= \left[a_\lambda^\dagger(\boldsymbol{q}), a_{\lambda'}^\dagger(\boldsymbol{q}')\right] = 0\,. \end{aligned} \qquad (12.1.33)$$

The inverse formulas are then

$$\begin{aligned} Q_\lambda(\boldsymbol{q}) &= \sqrt{\frac{\hbar}{2\omega_\lambda(\boldsymbol{q})}}\left[a_\lambda(\boldsymbol{q}) + a_\lambda^\dagger(-\boldsymbol{q})\right], \\ P_\lambda(\boldsymbol{q}) &= -\mathrm{i}\sqrt{\frac{\hbar\omega_\lambda(\boldsymbol{q})}{2}}\left[a_\lambda(-\boldsymbol{q}) - a_\lambda^\dagger(\boldsymbol{q})\right]. \end{aligned} \qquad (12.1.34)$$

In terms of these operators, the Hamiltonian is

$$\mathcal{H} = \tfrac{1}{2}\sum_{\boldsymbol{q},\lambda}\hbar\omega_\lambda(\boldsymbol{q})\left[a_\lambda^\dagger(\boldsymbol{q})a_\lambda(\boldsymbol{q}) + a_\lambda(-\boldsymbol{q})a_\lambda^\dagger(-\boldsymbol{q})\right]. \qquad (12.1.35)$$

Alternatively, by exploiting that $\omega_\lambda(\boldsymbol{q})$ is an even function of $\boldsymbol{q}$,

$$\boxed{\mathcal{H} = \sum_{\boldsymbol{q},\lambda}\hbar\omega_\lambda(\boldsymbol{q})\left[a_\lambda^\dagger(\boldsymbol{q})a_\lambda(\boldsymbol{q}) + \tfrac{1}{2}\right].} \qquad (12.1.36)$$

As the eigenvalues of operator $\hat{n}_\lambda(\boldsymbol{q}) = a_\lambda^\dagger(\boldsymbol{q})a_\lambda(\boldsymbol{q})$ are nonnegative integers, the vibrations of the crystal lattice are quantized in such a way that their energy can change by integral multiples of $\hbar\omega_\lambda(\boldsymbol{q})$. Operator $a_\lambda^\dagger(\boldsymbol{q})$ takes an initial state into one whose energy is higher by $\hbar\omega_\lambda(\boldsymbol{q})$ – while operator $a_\lambda(\boldsymbol{q})$

into another whose energy is lower by the same amount. These operators can be regarded as the creation and annihilation operators of a vibrational quantum.

Making use of the above form of the Hamiltonian, the time dependence of the creation and annihilation operators is

$$\begin{aligned} a_\lambda(\boldsymbol{q},t) =& \mathrm{e}^{\mathrm{i}\mathcal{H}t/\hbar} a_\lambda(\boldsymbol{q}) \mathrm{e}^{-\mathrm{i}\mathcal{H}t/\hbar} = a_\lambda(\boldsymbol{q}) \mathrm{e}^{-\mathrm{i}\omega_\lambda(\boldsymbol{q})t}\,, \\ a_\lambda^\dagger(\boldsymbol{q},t) =& \mathrm{e}^{\mathrm{i}\mathcal{H}t/\hbar} a_\lambda^\dagger(\boldsymbol{q}) \mathrm{e}^{-\mathrm{i}\mathcal{H}t/\hbar} = a_\lambda^\dagger(\boldsymbol{q}) \mathrm{e}^{\mathrm{i}\omega_\lambda(\boldsymbol{q})t}. \end{aligned} \tag{12.1.37}$$

Using them in the time derivative of equations (12.1.34) for $Q_\lambda(\boldsymbol{q})$ and $P_\lambda(\boldsymbol{q})$, it can be directly shown that (11.3.28), (11.3.29), and (11.3.31) are indeed satisfied.

Substituting (12.1.34) in the formula for atomic displacements, the displacement operator (11.3.14) may also be expressed in terms of these creation and annihilation operators:

$$\begin{aligned} u_\alpha(m,\mu) &= \sum_{\boldsymbol{q},\lambda} \sqrt{\frac{\hbar}{2NM_\mu\omega_\lambda(\boldsymbol{q})}} e_{\mu,\alpha}^{(\lambda)}(\boldsymbol{q}) \left[a_\lambda(\boldsymbol{q}) + a_\lambda^\dagger(-\boldsymbol{q})\right] \mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{R}_m} \\ &= \sum_{\boldsymbol{q},\lambda} \sqrt{\frac{\hbar}{2NM_\mu\omega_\lambda(\boldsymbol{q})}} \left[e_{\mu,\alpha}^{(\lambda)}(\boldsymbol{q}) a_\lambda(\boldsymbol{q}) \mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{R}_m} e_{\mu,\alpha}^{(\lambda)\,*}(\boldsymbol{q}) a_\lambda^\dagger(\boldsymbol{q}) \mathrm{e}^{-\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{R}_m}\right]. \end{aligned} \tag{12.1.38}$$

Writing out the time dependence explicitly,

$$\begin{aligned} u_\alpha(m,\mu,t) = \sum_{\boldsymbol{q},\lambda} \sqrt{\frac{\hbar}{2NM_\mu\omega_\lambda(\boldsymbol{q})}} \Big[& e_{\mu,\alpha}^{(\lambda)}(\boldsymbol{q}) a_\lambda(\boldsymbol{q}) \mathrm{e}^{\mathrm{i}(\boldsymbol{q}\cdot\boldsymbol{R}_m - \omega_\lambda(\boldsymbol{q})t)} \\ &+ e_{\mu,\alpha}^{(\lambda)\,*}(\boldsymbol{q}) a_\lambda^\dagger(\boldsymbol{q}) \mathrm{e}^{-\mathrm{i}(\boldsymbol{q}\cdot\boldsymbol{R}_m - \omega_\lambda(\boldsymbol{q})t)}\Big]. \end{aligned} \tag{12.1.39}$$

12.1.6 Phonons as Elementary Excitations

Apart from the term due to zero-point vibrations, the above form of the Hamiltonian of a thermally vibrating lattice is the same as the Hamiltonian of a free gas of bosonic particles with energy $\hbar\omega_\lambda(\boldsymbol{q})$. The same formula is obtained when the Hamiltonian of the radiation field is expressed in terms of the creation and annihilation operators of photons, the quanta of the electromagnetic field. Since acoustic elastic waves are related to sound propagation, the quanta of lattice vibrations are called *phonons* by this analogy, regardless of whether they are obtained through the quantization of acoustic or optical classical lattice vibrations. Consequently one may distinguish acoustic and optical phonons, and for both types longitudinal and transverse modes. The dispersion relation of acoustic phonons always starts linearly at $\boldsymbol{q} = 0$, while

that of optical phonons starts at a nonvanishing value, and generally varies only slightly.

The thermally excited collective vibrational states of a lattice can be regarded as states of a gas of fictitious particles – phonons – propagating in the lattice, independently of each other. States of the system can be characterized by the number of excited phonons. Since phonon creation and annihilation operators satisfy bosonic commutation relations, the *Bose–Einstein statistics*[2] applies to these fictitious particles. The thermal population of phonon states, i.e., the expectation value of the particle number operator is given by

$$\left\langle n_\lambda(\boldsymbol{q})\right\rangle \equiv \left\langle a_\lambda^\dagger(\boldsymbol{q})a_\lambda(\boldsymbol{q})\right\rangle = \frac{1}{\mathrm{e}^{\hbar\omega_\lambda(\boldsymbol{q})/k_\mathrm{B}T}-1}\,. \tag{12.1.40}$$

As the number of phonons is not conserved, the chemical potential is zero.

Starting with the ground state that has only zero-point vibrations, each subsequent action of the operators $a_\lambda^\dagger(\boldsymbol{q})$ increases the number of phonons by one and the energy by $\hbar\omega_\lambda(\boldsymbol{q})$. Just like in the classical model, frequencies are determined by the eigenvalues of the dynamical matrix. Consequently, $a_\lambda^\dagger(\boldsymbol{q})$ is the creation operator of a phonon of wave vector $\boldsymbol{q}$ and polarization λ. Similarly, the annihilation operator $a_\lambda(\boldsymbol{q})$ reduces the number of phonons by one and the energy of the state by $\hbar\omega_\lambda(\boldsymbol{q})$.

The energy contribution of lattice vibrations is calculated as the expectation value of the Hamiltonian (12.1.36),

$$E = \langle\mathcal{H}\rangle = \sum_{\boldsymbol{q},\lambda}\hbar\omega_\lambda(\boldsymbol{q})\left[\left\langle a_\lambda^\dagger(\boldsymbol{q})a_\lambda(\boldsymbol{q})\right\rangle + \tfrac{1}{2}\right]. \tag{12.1.41}$$

Using Bose–Einstein statistics,

$$E = \sum_{\boldsymbol{q},\lambda}\hbar\omega_\lambda(\boldsymbol{q})\left[\frac{1}{\mathrm{e}^{\hbar\omega_\lambda(\boldsymbol{q})/k_\mathrm{B}T}-1} + \tfrac{1}{2}\right]. \tag{12.1.42}$$

This essentially confirms the result obtained via naive quantization. An important difference with those formulas is the presence of the term 1/2, which corresponds to zero-point vibrations. We shall discuss its significance later.

Due care must be exercised in the calculation of the momentum of a vibrating lattice. Starting with expression (12.1.39), and considering the contribution of a phonon of specific wave vector and polarization, the momentum associated with the vibration is

$$\sum_{m,\mu}M_\mu\dot{u}_\alpha(m,\mu) = -\mathrm{i}\sum_{m,\mu}\sqrt{\frac{\hbar M_\mu\omega_\lambda(\boldsymbol{q})}{2N}}\left[e_{\mu,\alpha}^{(\lambda)}(\boldsymbol{q})a_\lambda(\boldsymbol{q})\mathrm{e}^{\mathrm{i}(\boldsymbol{q}\cdot\boldsymbol{R}_m-\omega_\lambda(\boldsymbol{q})t)}\right.$$
$$\left. - e_{\mu,\alpha}^{(\lambda)\,*}(\boldsymbol{q})a_\lambda^\dagger(\boldsymbol{q})\mathrm{e}^{-\mathrm{i}(\boldsymbol{q}\cdot\boldsymbol{R}_m-\omega_\lambda(\boldsymbol{q})t)}\right]. \tag{12.1.43}$$

[2] S. N. Bose, 1924; A. Einstein, 1924.

Summation over m, the label of primitive cells gives a vanishing result unless $\boldsymbol{q} = 0$. Nevertheless $\hbar\boldsymbol{q}$ – or sometimes $\boldsymbol{q}$ itself – is called the crystal momentum of the phonon: as discussed in Chapter 6, in processes where the phonon number changes by one, a conservation law – that is valid only up to a reciprocal-lattice vector – applies to the wave vector $\boldsymbol{q}$.

12.1.7 Acoustic Phonons as Goldstone Bosons

The Hamiltonian of a crystalline structure is invariant under arbitrary translations, while the crystal itself only under discrete ones. The crystalline state therefore breaks a continuous symmetry – or, more precisely, one along each spatial direction. According to Goldstone's theorem (page 200) when a continuous symmetry is broken, boson-like low-energy excitations, so-called soft modes or *Goldstone bosons* appear. The three acoustic branches starting at the center of the Brillouin zone at zero energy have precisely these properties. In this sense acoustic phonons can be considered Goldstone bosons.

The appearance of soft Goldstone bosons may be illustrated by a simple physical picture: in the acoustic branches excitation energies vanish in the long-wavelength limit, as the uniform translation of the lattice requires no expense of energy.

It should be noted that Goldstone's theorem does not apply for long-range interatomic forces. This can be directly demonstrated in a straightforward fashion by considering the long-wavelength limit of the frequency formula (11.2.14) obtained for monatomic chains. For small values of q the expansion yields

$$\omega = \sqrt{\frac{1}{M}\sum_p K_p\, p^2}\, qa\,. \tag{12.1.44}$$

When K_p decreases slowly with distance, the above sum diverges, indicating that the frequency does not vanish in the $q \to 0$ limit: it tends to a finite value.

12.1.8 Symmetries of the Vibrational Spectrum

It was shown in Chapter 6 that if a crystal is taken into itself by an element $P = \{\alpha|\boldsymbol{t}_n\}$ of the space group, then the same energy eigenvalue is associated with $\boldsymbol{q}$ and $\alpha\boldsymbol{q}$. In our case this means that the energy spectrum of phonons possesses the symmetry

$$\hbar\omega_\lambda(\boldsymbol{q}) = \hbar\omega_\lambda(\alpha\boldsymbol{q})\,. \tag{12.1.45}$$

This does not imply that the dynamical matrix itself should be invariant under the symmetry operations associated with the elements of the space group. It can nevertheless be shown that there exists a matrix $\mathsf{R}_P(\boldsymbol{q})$ such that the dynamical matrix satisfies the congruence relation

$$\mathsf{R}_P(\boldsymbol{q})\mathsf{D}(\boldsymbol{q})\mathsf{R}_P^{-1}(\boldsymbol{q}) = \mathsf{D}(\alpha\boldsymbol{q})\,. \tag{12.1.46}$$

By constructing a matrix $\mathsf{E}(\boldsymbol{q})$ from the eigenvectors as column vectors, the eigenvalue problem of the dynamical matrix can be written as

$$\mathsf{D}(\boldsymbol{q})\mathsf{E}(\boldsymbol{q}) = \omega^2(\boldsymbol{q})\mathsf{E}(\boldsymbol{q})\,, \tag{12.1.47}$$

where $\omega^2(\boldsymbol{q})$ is a diagonal matrix, and its diagonal elements are the eigenvalues. Applying matrix $\mathsf{R}_P(\boldsymbol{q})$ to this equation,

$$\mathsf{R}_P(\boldsymbol{q})\mathsf{D}(\boldsymbol{q})\mathsf{E}(\boldsymbol{q}) = \omega^2(\boldsymbol{q})\mathsf{R}_P(\boldsymbol{q})\mathsf{E}(\boldsymbol{q})\,. \tag{12.1.48}$$

On the other hand, from (12.1.46) we have

$$\mathsf{R}_P(\boldsymbol{q})\mathsf{D}(\boldsymbol{q})\mathsf{E}(\boldsymbol{q}) = \mathsf{D}(\alpha\boldsymbol{q})\mathsf{R}_P(\boldsymbol{q})\mathsf{E}(\boldsymbol{q})\,. \tag{12.1.49}$$

By comparing the two equations it is readily seen that the same frequencies are associated with $\alpha\boldsymbol{q}$ and $\boldsymbol{q}$; moreover, the relations among eigenvectors can also be established.

12.2 Density of Phonon States

When determining macroscopic quantities one frequently faces the problem of summing the contributions of individual phonons over all allowed vectors $\boldsymbol{q}$ in the Brillouin zone. If the crystal contains N_1, N_2, and N_3 primitive cells in the directions of the three primitive translation vectors then, according to (6.2.22), the vectors $\boldsymbol{q}$ allowed by the periodic boundary conditions are separated by $\boldsymbol{b}_1/N_1$, $\boldsymbol{b}_2/N_2$, and $\boldsymbol{b}_3/N_3$ along the directions of the three primitive vectors of the reciprocal lattice. In samples that are large compared to atomic dimensions the vectors $\boldsymbol{q}$ are spaced rather densely, therefore no significant error arises from replacing summation by integration.

Since the volume associated with each allowed point $\boldsymbol{q}$ of the reciprocal lattice is

$$\Delta\boldsymbol{q} = \frac{\boldsymbol{b}_1}{N_1}\cdot\left(\frac{\boldsymbol{b}_2}{N_2}\times\frac{\boldsymbol{b}_3}{N_3}\right) = \frac{1}{N}[\boldsymbol{b}_1\boldsymbol{b}_2\boldsymbol{b}_3] = \frac{v_r}{N} = \frac{1}{N}\frac{(2\pi)^3}{v} = \frac{(2\pi)^3}{V}\,, \tag{12.2.1}$$

where $N = N_1\,N_2\,N_3$, and v_r is the volume of the primitive cell of the reciprocal lattice, while v is that of the direct lattice, replacing the sum by an integral requires the following substitution:

$$\boxed{\sum_{\boldsymbol{q}} \to \frac{V}{(2\pi)^3}\int \mathrm{d}\boldsymbol{q}.} \tag{12.2.2}$$

If the summand (integrand) is a function of the phonon energy alone, then it is usually much simpler to calculate the integral using the density of states (DOS), i.e., the number of phonons of a given energy.

12.2.1 Definition of the Density of States

The total density of states $g(\omega)$ and the partial density of states $g_\lambda(\omega)$ of a given polarization branch λ is defined through the requirement that for any function $f(\omega)$

$$\begin{aligned}\frac{1}{V}\sum_{\boldsymbol{q},\lambda} f(\omega_\lambda(\boldsymbol{q})) &= \sum_\lambda \int \frac{\mathrm{d}\boldsymbol{q}}{(2\pi)^3} f(\omega_\lambda(\boldsymbol{q})) \\ &= \sum_\lambda \int \mathrm{d}\omega\, g_\lambda(\omega) f(\omega) = \int \mathrm{d}\omega\, g(\omega) f(\omega)\,.\end{aligned} \tag{12.2.3}$$

Then the partial and total densities of states are formally given by

$$\begin{aligned} g_\lambda(\omega) &= \int \frac{\mathrm{d}\boldsymbol{q}}{(2\pi)^3}\delta(\omega-\omega_\lambda(\boldsymbol{q}))\,, \\ g(\omega) &= \sum_\lambda \int \frac{\mathrm{d}\boldsymbol{q}}{(2\pi)^3}\delta(\omega-\omega_\lambda(\boldsymbol{q}))\,. \end{aligned} \tag{12.2.4}$$

In the general d-dimensional case $(2\pi)^3$ is replaced by $(2\pi)^d$.[3]

In the Einstein model the frequency ω_E is the same for each vibration, so

$$g_\mathrm{E}(\omega) = \int \frac{\mathrm{d}\boldsymbol{q}}{(2\pi)^3}\delta(\omega-\omega_\mathrm{E}) = \frac{N}{V}\delta(\omega-\omega_\mathrm{E})\,. \tag{12.2.5}$$

In the second step we made use of relation (5.2.18) between the volumes of the Brillouin zone and the primitive cell. This delta-like sharp peak in the density of states is often a good approximation for optical phonons, since their group velocity is low.

To evaluate the density of states numerically from the energy spectrum, it should be remembered that $g_\lambda(\omega)\,\mathrm{d}\omega$ is the number of states with energies between $\hbar\omega$ and $\hbar(\omega+\mathrm{d}\omega)$ divided by V. It is sometimes more convenient to determine the density of states from $\mathcal{N}_\lambda(\omega)$, which is the number of vibrational states with polarization λ and frequency less than ω in the total volume of the sample. The two quantities are related by

$$g_\lambda(\omega) = \frac{1}{V}\frac{\mathrm{d}\mathcal{N}_\lambda(\omega)}{\mathrm{d}\omega}\,, \tag{12.2.6}$$

which is why $\mathcal{N}_\lambda(\omega)$ is sometimes called the *integrated density of states* (IDS).

To determine the density of states, consider the surfaces of constant energy $S_\lambda(\omega)$ and $S_\lambda(\omega+\mathrm{d}\omega)$ in reciprocal space, shown in Fig. 12.2.

[3] As the eigenvalues of the dynamical matrix are given as functions of ω^2, the density of states is often written in terms of $D(\omega^2) = g(\omega)/(2\omega)$ rather than $g(\omega)$. $D(\omega^2)\mathrm{d}\omega^2$ is the number of those vibrational states whose squared frequency is between ω^2 and $\omega^2+\mathrm{d}\omega^2$.

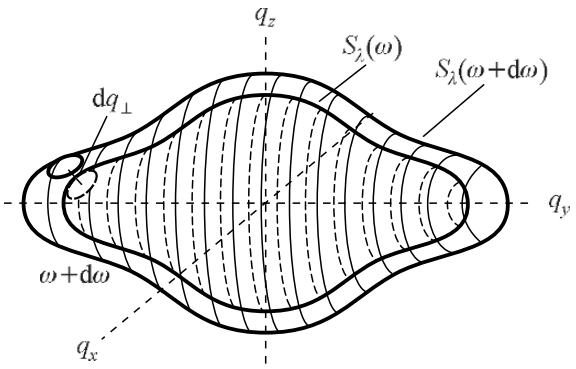


Fig. 12.2. The surfaces of constant energy $\hbar\omega$ and $\hbar(\omega+\mathrm{d}\omega)$. The density of states is given by the number of allowed wave vectors between the two surfaces

Apart from a factor $1/V$, $g_\lambda(\omega)\mathrm{d}\omega$ is, by definition, the number of allowed vectors $\boldsymbol{q}$ in the region between the two surfaces. This is obtained by dividing the volume of this region by the volume $(2\pi)^3/V$ associated with the vector $\boldsymbol{q}$. Denoting the perpendicular distance of the two surfaces by $\mathrm{d}q_\perp(\boldsymbol{q})$ and the surface element by $\mathrm{d}S$ for any vector $\boldsymbol{q}$ that satisfies the condition $\omega_\lambda(\boldsymbol{q})=\omega$, the volume of the region between the two surfaces is

$$\int\limits_{S_\lambda(\omega)} \mathrm{d}S\mathrm{d}q_\perp(\boldsymbol{q})\,, \tag{12.2.7}$$

and so the number of states inside the region is

$$g_\lambda(\omega)\,\mathrm{d}\omega = \int\limits_{S_\lambda(\omega)} \frac{\mathrm{d}S}{(2\pi)^3}\mathrm{d}q_\perp(\boldsymbol{q})\,. \tag{12.2.8}$$

Using a linear expansion for the dispersion relation,

$$\omega+\mathrm{d}\omega = \omega + |\boldsymbol{\nabla}_{\boldsymbol{q}}\omega_\lambda(\boldsymbol{q})|\mathrm{d}q_\perp(\boldsymbol{q})\,, \tag{12.2.9}$$

from which

$$\mathrm{d}q_\perp(\boldsymbol{q}) = \frac{\mathrm{d}\omega}{|\boldsymbol{\nabla}_{\boldsymbol{q}}\omega_\lambda(\boldsymbol{q})|}\,. \tag{12.2.10}$$

Substituting this into (12.2.8) leads to the following formula for the density of states:

$$\boxed{g_\lambda(\omega) = \frac{1}{(2\pi)^3}\int\limits_{S_\lambda(\omega)} \frac{\mathrm{d}S}{|\boldsymbol{\nabla}_{\boldsymbol{q}}\omega_\lambda(\boldsymbol{q})|}\,.} \tag{12.2.11}$$

If the dispersion relation is approximated by that of elastic waves,

$$\omega_\lambda(\boldsymbol{q}) = c_\lambda|\boldsymbol{q}| \tag{12.2.12}$$

as in the Debye model, but with polarization-dependent propagation velocity, then the size of the constant-energy surface of energy $\hbar\omega$ is

$$S_\lambda(\omega) = 4\pi \left(\frac{\omega}{c_\lambda}\right)^2 . \tag{12.2.13}$$

The partial density of states for phonons of polarization λ is then

$$g_\lambda(\omega) = \frac{1}{(2\pi)^3} 4\pi \left(\frac{\omega}{c_\lambda}\right)^2 \frac{1}{c_\lambda} = \frac{1}{2\pi^2} \frac{\omega^2}{c_\lambda^3} . \tag{12.2.14}$$

Assuming that the propagation velocity is the same for both transverse vibrations, the total density of states is

$$g(\omega) = \frac{1}{2\pi^2} \omega^2 \left[\frac{1}{c_\mathrm{L}^3} + \frac{2}{c_\mathrm{T}^3}\right] . \tag{12.2.15}$$

As it has been mentioned, the same velocity c_D is associated with each of the three vibrational branches in the Debye model. Naturally, this value has to be chosen in such a way that the correct density of states is recovered for long-wavelength phonons. It follows directly from the above formula that

$$\frac{3}{c_\mathrm{D}^3} = \frac{1}{c_\mathrm{L}^3} + \frac{2}{c_\mathrm{T}^3} . \tag{12.2.16}$$

However, the ensuing formula for the density of states,

$$g_\mathrm{D}(\omega) = \frac{3}{2\pi^2} \frac{\omega^2}{c_\mathrm{D}^3} \tag{12.2.17}$$

cannot be correct for very large energies: there is a maximum wave number q_D in the Debye model, determined by (12.1.13), and therefore vibrational frequencies cannot exceed the *Debye frequency*

$$\omega_\mathrm{D} = c_\mathrm{D}\, q_\mathrm{D} . \tag{12.2.18}$$

Expressed in terms of this quantity, the density of states is

$$g_\mathrm{D}(\omega) = \begin{cases} 9p \dfrac{N}{V} \dfrac{\omega^2}{\omega_\mathrm{D}^3} , & \text{if} \quad \omega \le \omega_\mathrm{D} , \\ 0 , & \text{if} \quad \omega > \omega_\mathrm{D} . \end{cases} \tag{12.2.19}$$

This form of the density of states is the consequence of the linearity of the dispersion relation and the dimensionality (3) of the direct and reciprocal space. The quadratic increase of the density of states is true quite generally, even beyond the Debye model – but only for small values of ω, where the dispersion relation of acoustic phonons is nearly linear. In the general case of a d-dimensional crystal and in the same limit (i.e., for small values of ω) the density of states is proportional to the $d-1$st power of frequency,

$$g(\omega) \propto \omega^{d-1} . \tag{12.2.20}$$

12.2.2 The Density of States in One- and Two-Dimensional Systems

Before turning to the discussion of the singularities in the density of states for three-dimensional crystals it is instructive to study two simpler cases: those of one-dimensional chains and of two-dimensional lattices.

In the one-dimensional case the allowed values of q are spaced $2\pi/L$ apart. Since the same frequency is associated with q and $-q$, the number of states in $\mathrm{d}\omega$ is

$$\frac{L}{\pi}\frac{\mathrm{d}q}{\mathrm{d}\omega}\mathrm{d}\omega\,. \tag{12.2.21}$$

The density of states, that is, the number of states with frequency ω per unit length of the chain is then

$$g_{1d}(\omega) = \frac{1}{\pi}\frac{1}{\mathrm{d}\omega/\mathrm{d}q}\,. \tag{12.2.22}$$

When writing the dispersion relation (11.2.9) for monatomic chains in the form $\omega(q) = \omega_{\max}|\sin(qa/2)|$, the following relation emerges:

$$g_{1d}(\omega) = \begin{cases} \dfrac{2}{\pi a}\left(\omega_{\max}^2 - \omega^2\right)^{-1/2}, & \text{if} \quad \omega \le \omega_{\max}\,, \\ 0\,, & \text{if} \quad \omega > \omega_{\max}\,. \end{cases} \tag{12.2.23}$$

The density of states is seen to be singular for phonon energies at the Brillouin zone boundary. This inverse-square-root singularity is characteristic of the one-dimensional case. Similar singularities are observed in the density of states of the optical vibrations in diatomic chains. However, in the latter case the inverse-square-root singularities occur not only for frequency maxima but for frequency minima as well. In the acoustic branch the density of states does not show any singularity around $\omega = 0$, as for small values of q the dispersion relation is not quadratic but linear. This apparently nonanalytic behavior – $\omega \propto |q|$ – is the consequence of the fact that the eigenvalues of the dynamical matrix lead to analytical expressions in q for ω^2 (rather than ω).

In two-dimensional crystals, too, each phonon branch contains at least one point $\boldsymbol{q}_0$ where the dispersion relation attains a maximum. In the vicinity of this point the energy spectrum is described by a quadratic form of the components of the vector $\boldsymbol{\xi} = \boldsymbol{q} - \boldsymbol{q}_0$. The principal axis transformation of the quadratic expression gives

$$\omega_\lambda(\boldsymbol{q}) = \omega_0 - \alpha_1\xi_1^2 - \alpha_2\xi_2^2\,. \tag{12.2.24}$$

Using the variables $x_i = \alpha_i^{1/2}\xi_i$, the density of states can be determined with the help of the integral

$$g_\lambda(\omega) = \frac{1}{(2\pi)^2}\frac{1}{(\alpha_1\alpha_2)^{1/2}}\iint \mathrm{d}x_1\,\mathrm{d}x_2\,\delta(\omega - \omega_0 + x_1^2 + x_2^2)\,. \tag{12.2.25}$$

By changing to polar coordinates,

$$g_\lambda(\omega) = \frac{1}{(2\pi)^2}\frac{1}{(\alpha_1\alpha_2)^{1/2}}\iint r\,\mathrm{d}r\,\mathrm{d}\varphi\,\delta(\omega-\omega_0+r^2)$$
$$= \begin{cases} \dfrac{1}{4\pi}\dfrac{1}{(\alpha_1\alpha_2)^{1/2}}, & \text{if} \quad \omega \le \omega_0\,, \\ 0\,, & \text{if} \quad \omega > \omega_0\,. \end{cases} \tag{12.2.26}$$

This means that the density of states drops to zero from a finite value at the maximum frequency. A similar method can be applied to the case when the dispersion curve has a minimum around $\boldsymbol{q}_0$; the density of states then jumps from zero to a finite value at the corresponding energy. In line with the general considerations presented above, the dispersion relation starts linearly at $\boldsymbol{q} = 0$ in the acoustic branches, and therefore the density of states starts linearly at the bottom of the spectrum.

In each phonon branch there is at least one so-called saddle point where the dispersion relation attains a minimum in one direction and a maximum in the other. This is easily demonstrated by allowing equivalent wave vectors and considering the $\boldsymbol{q}$ over the entire reciprocal space. Then, without loss of generality, the Brillouin zone can be centered at the point $\boldsymbol{q}_0$ where the dispersion curve attains a minimum. At any equivalent point $\boldsymbol{q}_0+\boldsymbol{G}$, where $\boldsymbol{G}$ is a primitive vector of the reciprocal lattice, the dispersion relation attains another minimum. When points $\boldsymbol{q}_0$ and $\boldsymbol{q}_0+\boldsymbol{G}$ are connected through various paths in $\boldsymbol{q}$-space, there exists a local maximum along each path. These paths and the local maxima along them are shown in Fig. 12.3. Next the curve connecting these maxima is considered: its points at the boundaries of the Brillouin zone are equivalent, as they differ by a reciprocal-lattice vector. Consequently there must be a local minimum along this curve – which is a saddle point.

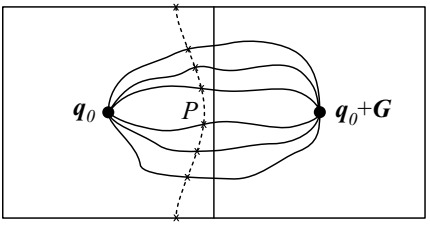


Fig. 12.3. Determination of the locus of the saddle point in the extended reciprocal space. Crosses (×) mark the maxima of the dispersion curve along the paths connecting $\boldsymbol{q}_0$, the center of the Brillouin zone, with an equivalent point $\boldsymbol{q}_0+\boldsymbol{G}$. The minimum along the curve connecting these maxima is the saddle point P

In the vicinity of the saddle point the dispersion relation takes the form

$$\omega_\lambda(\boldsymbol{q}) = \omega_0 + \alpha_1\xi_1^2 - \alpha_2\xi_2^2 \tag{12.2.27}$$

in the system of principal axes, with positive coefficients α_i. As shown in Fig. 12.4, the lines of constant energy are hyperbolas.

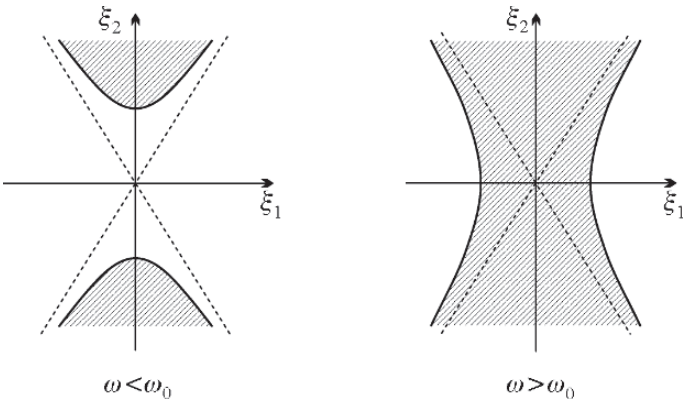


Fig. 12.4. Lines of constant energy around the saddle point for $\omega < \omega_0$ and $\omega > \omega_0$. Phonon energies are lower in the shaded regions than on the curves bounding them

Once again, we introduce the variables $x_i = \alpha_i^{1/2}\xi_i$; the density of states is then given by

$$g_\lambda(\omega) = \frac{1}{(2\pi)^2}\frac{1}{(\alpha_1\alpha_2)^{1/2}}\iint \mathrm{d}x_1\,\mathrm{d}x_2\,\delta(\omega-\omega_0-x_1^2+x_2^2)\,. \qquad (12.2.28)$$

To evaluate the integral, we change to coordinates that are adapted to the hyperbola. A good choice is

$$x_1 = r\sinh\theta\,, \qquad x_2 = \pm r\cosh\theta \qquad (12.2.29)$$

when $\omega < \omega_0$, and

$$x_1 = \pm r\cosh\theta\,, \qquad x_2 = r\sinh\theta \qquad (12.2.30)$$

when $\omega > \omega_0$. Using the Jacobian of the new variables,

$$g_\lambda(\omega) = \frac{1}{4\pi^2}\frac{1}{(\alpha_1\alpha_2)^{1/2}}\iint r\,\mathrm{d}r\,\mathrm{d}\theta\,\delta(\omega-\omega_0\pm r^2)\,. \qquad (12.2.31)$$

The sign $(\pm)$ in front of r^2 corresponds to the cases $\omega < \omega_0$ and $\omega > \omega_0$. In either case, the contribution vanishes unless $r = |\omega-\omega_0|^{1/2}$. However, the value of the integral depends on the choice of the range of integration for the variable θ. Assuming that the quadratic approximation for the dispersion curve is valid in a finite neighborhood of the saddle point, integration is performed in the region where

$$x_1^2 + x_2^2 \le R^2\,. \qquad (12.2.32)$$

Even though the cutoff R is introduced arbitrarily, the physically meaningful singularity does not depend on its particular choice.

Since the two branches of the hyperbola give equal contributions, and so do the two arms of each branch, integration needs to be performed only in one quadrant, e.g., $x_1 > 0$, $x_2 > 0$. To include only points within the circle of radius R, the following restriction must be imposed on the variable θ:

$$0 \leq \theta \leq \operatorname{arsinh}\left[\tfrac{1}{2}\left(\frac{R^2}{r^2} - 1\right)\right]^{1/2}. \tag{12.2.33}$$

Integration then gives

$$g_\lambda(\omega) = \frac{1}{2\pi^2}\frac{1}{(\alpha_1\alpha_2)^{1/2}} \operatorname{arsinh}\left[\tfrac{1}{2}\left(\frac{R^2}{|\omega_0 - \omega|} - 1\right)\right]^{1/2}. \tag{12.2.34}$$

Using the logarithmic expression

$$\operatorname{arsinh} x = \ln[x + (x^2 + 1)^{1/2}] \tag{12.2.35}$$

for the inverse hyperbolic function, $\operatorname{arsinh} x \approx \ln 2x$ in the vicinity of the saddle point, where the argument is large. Therefore a logarithmic singularity appears in the density of states:

$$g_\lambda(\omega) \approx -\frac{1}{4\pi^2}\frac{1}{(\alpha_1\alpha_2)^{1/2}} \ln\left|1 - \frac{\omega}{\omega_0}\right|. \tag{12.2.36}$$

12.2.3 Van Hove Singularities

In the one-dimensional case the density of states has inverse-square-root singularities at frequencies where the spectrum has maxima or minima. In two dimensions, in addition to maxima and minima, the spectrum may also have saddle points; the density of states then features finite jumps and logarithmic singularities, respectively. In the Debye model of three-dimensional crystals the spectrum is assumed to be isotropic, therefore the density of states is smooth everywhere except for an abrupt cutoff at the maximum energy. This is not the case for real crystals. Energy extrema may occur at several points of the phonon spectrum, and then for the corresponding $\boldsymbol{q}_0$ value(s) $|\boldsymbol{\nabla}_{\boldsymbol{q}}\,\omega_\lambda(\boldsymbol{q})| = 0$ along every direction. In other cases the gradient of the dispersion relation vanishes in certain directions only. As we saw in Chapter 6, it always vanishes at the boundaries of the Brillouin zone in the perpendicular direction if the boundary is related to the opposite one through reflection symmetry. At the corners of the Brillouin zone and at the centers of the boundary faces the gradients in other directions may vanish as well. Those points where the gradient vanishes in every direction are called analytical critical points. At such points the integrand is singular in expression (12.2.11) for the density of states. Nevertheless the integral – and so the density of states – is finite

for any value of the frequency. The points where its derivative is singular are called *Van Hove singularities*.[4] Figure 12.5 shows the density of states derived from the theoretically determined phonon spectrum of aluminum. Singularities come from transverse branches at low energies and from longitudinal phonons at the top of the spectrum.

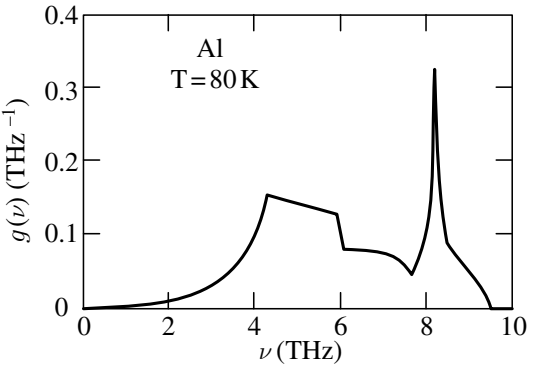


Fig. 12.5. Calculated phonon density of states in aluminum, with the characteristic Van Hove singularities [C. B. Walker, *Phys. Rev.* **103**, 547 (1956)]

When the phonon dispersion relation is expanded in a series of the components of $\boldsymbol{\xi} = \boldsymbol{q} - \boldsymbol{q}_0$ around an analytical critical point $\boldsymbol{q}_0$, the linear terms must be absent. The principal axis transformation of the second-order expression gives

$$\omega_\lambda(\boldsymbol{q}) = \omega_0 + \alpha_1\xi_1^2 + \alpha_2\xi_2^2 + \alpha_3\xi_3^2\,, \tag{12.2.37}$$

where $\omega_0 = \omega_\lambda(\boldsymbol{q}_0)$. In contrast to the minima, maxima, and saddle points of the two-dimensional case, four types of critical points are now distinguished.

In P_0-type points each coefficient α_i is positive, and the spectrum has a minimum at $\boldsymbol{q}_0$. We shall now show that when ω is close to the corresponding frequency, the density of states is

$$g_\lambda(\omega) = \frac{1}{4\pi^2}\frac{(\omega-\omega_0)^{1/2}}{(\alpha_1\alpha_2\alpha_3)^{1/2}}\,. \tag{12.2.38}$$

We shall first determine the number of states on the constant-energy surface associated with the frequency ω. This surface is an ellipsoid that intersects the ξ_i-axis in $\pm[(\omega-\omega_0)/\alpha_i]^{1/2}$. The volume of the ellipsoid is

$$\frac{4\pi}{3}\frac{(\omega-\omega_0)^{3/2}}{(\alpha_1\alpha_2\alpha_3)^{1/2}}\,. \tag{12.2.39}$$

The number of states within gives the integrated density of states:

[4] L. Van Hove, 1953.

$$\mathcal{N}_\lambda(\omega) = \frac{V}{(2\pi)^3}\frac{4\pi}{3}\frac{(\omega-\omega_0)^{3/2}}{(\alpha_1\alpha_2\alpha_3)^{1/2}} = \frac{V}{6\pi^2}\frac{(\omega-\omega_0)^{3/2}}{(\alpha_1\alpha_2\alpha_3)^{1/2}}\,. \tag{12.2.40}$$

When the derivative of this quantity with respect to frequency is divided by the volume, the density of states (12.2.38) is recovered.

In P_1-type points one of the α_i is negative and the two others are positive, and so the spectrum has a saddle point at $\boldsymbol{q}_0$. Surfaces of constant energy are hyperboloids of two sheets for $\omega < \omega_0$ and hyperboloids of one sheet for $\omega > \omega_0$. These are illustrated in Fig. 12.6. In the $\omega = \omega_0$ case the hyperboloid degenerates into a cone.

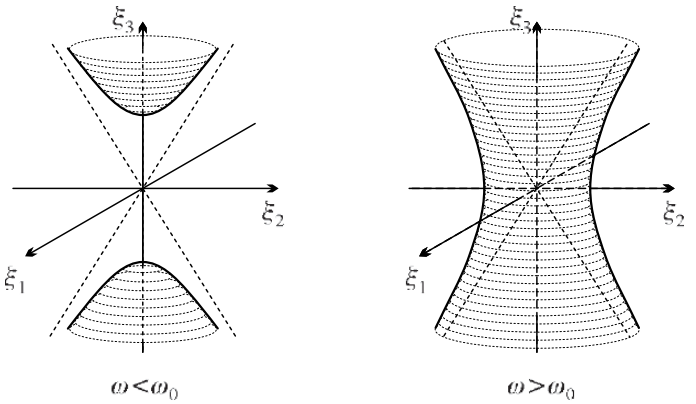


Fig. 12.6. Surfaces of constant energy around a P_1-type saddle point for $\omega < \omega_0$ and $\omega > \omega_0$

Using the form

$$\omega_\lambda(\boldsymbol{q}) = \omega_0 + \alpha_1\xi_1^2 + \alpha_2\xi_2^2 - \alpha_3\xi_3^2 \tag{12.2.41}$$

for describing the frequency spectrum, and changing to the variables $x_i = \alpha_i^{1/2}\xi_i$, the density of states can be expressed as

$$g_\lambda(\omega) = \frac{1}{(2\pi)^3}\frac{1}{(\alpha_1\alpha_2\alpha_3)^{1/2}}\iiint \mathrm{d}x_1\,\mathrm{d}x_2\,\mathrm{d}x_3\,\delta(\omega-\omega_0-x_1^2-x_2^2+x_3^2)\,. \tag{12.2.42}$$

To evaluate the integral, we shall introduce new coordinates that are adapted to the hyperboloid. We choose

$$x_1 = r\sinh\theta\cos\varphi, \qquad x_2 = r\sinh\theta\sin\varphi, \qquad x_3 = \pm r\cosh\theta \tag{12.2.43}$$

for $\omega < \omega_0$, and

$$x_1 = r\cosh\theta\cos\varphi, \qquad x_2 = r\cosh\theta\sin\varphi, \qquad x_3 = r\sinh\theta \tag{12.2.44}$$

for $\omega > \omega_0$. By evaluating the Jacobian for the new sets of variables, we have

$$g_\lambda(\omega) = \frac{1}{(2\pi)^3} \frac{1}{(\alpha_1\alpha_2\alpha_3)^{1/2}} \iiint r^2 \, \mathrm{d}r \sinh\theta \, \mathrm{d}\theta \, \mathrm{d}\varphi \, \delta(\omega - \omega_0 + r^2) \quad (12.2.45)$$

for $\omega < \omega_0$, and

$$g_\lambda(\omega) = \frac{1}{(2\pi)^3} \frac{1}{(\alpha_1\alpha_2\alpha_3)^{1/2}} \iiint r^2 \, \mathrm{d}r \cosh\theta \, \mathrm{d}\theta \, \mathrm{d}\varphi \, \delta(\omega - \omega_0 - r^2) \quad (12.2.46)$$

for $\omega > \omega_0$. In either case the integrand vanishes unless $r = |\omega - \omega_0|^{1/2}$. Assuming once again that the quadratic approximation for the dispersion curve is valid in a finite neighborhood of the saddle point, integration is performed in the region where

$$x_1^2 + x_2^2 + x_3^2 \le R^2 \, . \quad (12.2.47)$$

Positive and negative values of x_3 contribute equally, therefore it is sufficient to calculate the integral over positive values. To include only points within the sphere of radius R, the following restriction must be imposed on the variable θ:

$$\begin{aligned} 1 \le \cosh\theta &\le \left[\tfrac{1}{2}\left(\frac{R^2}{r^2} + 1\right)\right]^{1/2}, \qquad \text{if} \quad \omega < \omega_0 \, , \\ 0 \le \sinh\theta &\le \left[\tfrac{1}{2}\left(\frac{R^2}{r^2} - 1\right)\right]^{1/2}, \qquad \text{if} \quad \omega > \omega_0 \, . \end{aligned} \quad (12.2.48)$$

Integration with respect to θ is straightforward. Since the φ-integral gives a factor of 2π, we have

$$g_\lambda(\omega) = \frac{2}{(2\pi)^2} \frac{1}{(\alpha_1\alpha_2\alpha_3)^{1/2}} \int r^2 \, \mathrm{d}r \left\{ \left[\tfrac{1}{2}\left(\frac{R^2}{r^2} + 1\right)\right]^{1/2} - 1 \right\} \delta(\omega - \omega_0 + r^2) \quad (12.2.49)$$

for $\omega < \omega_0$, and

$$g_\lambda(\omega) = \frac{2}{(2\pi)^2} \frac{1}{(\alpha_1\alpha_2\alpha_3)^{1/2}} \int r^2 \, \mathrm{d}r \left[\tfrac{1}{2}\left(\frac{R^2}{r^2} - 1\right)\right]^{1/2} \delta(\omega - \omega_0 - r^2) \quad (12.2.50)$$

for $\omega > \omega_0$. Because of the delta functions these integrals are also elementary. For frequencies close to the saddle point ω_0, i.e., for $|\omega - \omega_0| \ll R^2$, the final result is

$$g_\lambda(\omega) = \begin{cases} C - \dfrac{1}{4\pi^2} \dfrac{(\omega_0 - \omega)^{1/2}}{|\alpha_1\alpha_2\alpha_3|^{1/2}} + \mathcal{O}(\omega_0 - \omega) \, , & \text{if} \quad \omega < \omega_0 \, , \\ C + \mathcal{O}(\omega - \omega_0) \, , & \text{if} \quad \omega > \omega_0 \, . \end{cases} \quad (12.2.51)$$

The density of states has a kink at ω_0, with a square-root singularity on one side.

In P_2-type analytical critical points two of the α_i are negative and the third is positive, so once again the spectrum has a saddle point at $\boldsymbol{q}_0$. The density of states exhibits a kink again, but the singularity is now on the other side:

$$g_\lambda(\omega) = \begin{cases} C + \mathcal{O}(\omega_0 - \omega)\,, & \text{if} \quad \omega < \omega_0\,, \\ C - \dfrac{1}{4\pi^2} \dfrac{(\omega - \omega_0)^{1/2}}{|\alpha_1 \alpha_2 \alpha_3|^{1/2}} + \mathcal{O}(\omega - \omega_0)\,, & \text{if} \quad \omega > \omega_0\,. \end{cases} \tag{12.2.52}$$

Finally, in P_3-type points each coefficient α_i is negative, and the spectrum has a minimum at $\boldsymbol{q}_0$. The density of states is

$$g_\lambda(\omega) = \frac{1}{4\pi^2} \frac{(\omega_0 - \omega)^{1/2}}{|\alpha_1 \alpha_2 \alpha_3|^{1/2}}. \tag{12.2.53}$$

Figure 12.7 shows the behavior of the density of states around the four kinds of the analytical critical point. It should be mentioned again that the velocity of acoustic phonons is finite at $\boldsymbol{q} = 0$, and so no Van Hove singularity appears at $\omega = 0$ in the phonon density of states. In three-dimensional systems the density of states is initially proportional to ω^2, in agreement with the Debye model.

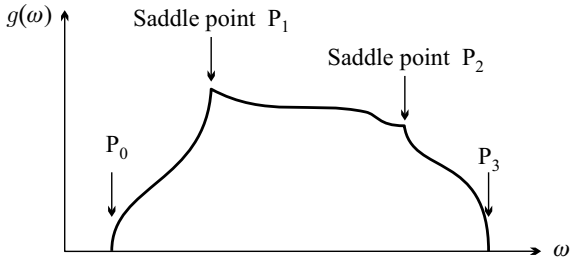


Fig. 12.7. Van Hove singularities in the phonon density of states

12.3 The Thermodynamics of Vibrating Lattices

From now on we consider the vibrating lattice as a gas of free phonons governed by the Bose–Einstein statistics, which can be characterized by the dispersion relation of the elementary excitations or the density of states. We shall examine how the thermodynamic properties of the crystal lattice may be interpreted in this picture.

12.3.1 The Ground State of the Lattice and Melting

In the ground state of a classical crystal atoms are fixed rigidly at their equilibrium positions. Such a state is ruled out in quantum mechanics: this is why the Hamiltonian contains, in addition to the particle number, a term 1/2, which corresponds to the energy of zero-point vibrations. If it were not for these zero-point vibrations, the mean square displacement of the atoms would vanish in the ground state.

The latter can be determined using (12.1.39), the formula for atomic displacements. Only those terms contribute to the average in which phonons of the same energy are created and annihilated, either in this or in the reverse order:

$$\begin{aligned}\left\langle u_\alpha^2(m,\mu)\right\rangle = \frac{\hbar}{2NM_\mu}\sum_{\boldsymbol{q},\lambda}\frac{1}{\omega_\lambda(\boldsymbol{q})}e_{\mu,\alpha}^{(\lambda)}(\boldsymbol{q})e_{\mu,\alpha}^{(\lambda)\,*}(\boldsymbol{q}) \\ \times\left[\left\langle a_\lambda(\boldsymbol{q})a_\lambda^\dagger(\boldsymbol{q})\right\rangle + \left\langle a_\lambda^\dagger(\boldsymbol{q})a_\lambda(\boldsymbol{q})\right\rangle\right].\end{aligned} \tag{12.3.1}$$

For simplicity, we shall consider a crystal with a monatomic basis, so the label μ will be absent. By summing over the three Cartesian coordinates, and by assuming that polarization vectors are normalized to unity, interchanging the order of the creation and annihilation operators yields

$$\left\langle \boldsymbol{u}^2(m)\right\rangle = \frac{\hbar}{NM}\sum_{\boldsymbol{q},\lambda}\frac{1}{\omega_\lambda(\boldsymbol{q})}\left[\left\langle a_\lambda^\dagger(\boldsymbol{q})a_\lambda(\boldsymbol{q})\right\rangle + \tfrac{1}{2}\right]. \tag{12.3.2}$$

Even though no phonons are present in the ground state, the mean square displacement of the atoms is nevertheless finite because of the quantum fluctuations:

$$\left\langle \boldsymbol{u}^2(m)\right\rangle = \frac{\hbar}{2NM}\sum_{\boldsymbol{q},\lambda}\frac{1}{\omega_\lambda(\boldsymbol{q})}, \tag{12.3.3}$$

which is the straightforward generalization of expression (12.1.22) for a single oscillator. Using the density of states, the left-hand side can be expressed as a frequency integral:

$$\left\langle \boldsymbol{u}^2(m)\right\rangle = \frac{\hbar V}{2NM}\int \mathrm{d}\omega\,\frac{g(\omega)}{\omega}. \tag{12.3.4}$$

According to (12.2.20), in d-dimensional crystals the density of states $g(\omega)$ is proportional to ω^{d-1} in the low-frequency limit, therefore in $d=2$ and $d=3$ dimensions integration up to the finite Debye frequency gives a finite result for the mean square displacement of the atoms. However, in $d=1$ dimension the integral diverges at the lower limit, indicating that one-dimensional ordered chains of atoms with discrete translational symmetry cannot exist. Quantum fluctuations are so strong that they destroy long-range order. This is in line

with Coleman's theorem presented in Chapter 6, which states that quantum fluctuations restore continuous symmetry in one dimension.

Obviously, the divergence has arisen because of the transformation of the discrete sum into an integral. Strictly speaking, this means that only in the thermodynamic limit, when the size of the sample approaches infinity, does the atomic displacement become infinitely large. In finite systems the wave vectors are finite and discrete, therefore the contribution of low-energy phonons is finite. This can be sufficiently large to destabilize regular, periodic arrays of atoms, nevertheless atoms may be arranged in one-dimensional chain-like patterns. This is observed in the compound $Hg_{3-\delta}AsF_6$, where mercury atoms occupy the tubes formed by AsF_6 molecules. As the arrangement is one-dimensional, the positions of the mercury atoms do not show long-range order, consequently no sharp Bragg peaks are observed in the diffraction pattern. However, there is a short-range order, since nearest-neighbor distances are roughly equal.

Expression (12.3.2) for the mean square atomic displacement is valid at finite temperatures as well, provided the average is taken as a thermal average. Using the Bose–Einstein statistics for the occupation number of the phonons,

$$\begin{aligned}\left\langle \boldsymbol{u}^2(m)\right\rangle &= \frac{\hbar}{NM}\sum_{\boldsymbol{q},\lambda}\frac{1}{\omega_\lambda(\boldsymbol{q})}\left[\frac{1}{\mathrm{e}^{\hbar\omega_\lambda(\boldsymbol{q})/k_\mathrm{B}T}-1}+\tfrac{1}{2}\right]\\ &= \frac{\hbar}{2NM}\sum_{\boldsymbol{q},\lambda}\frac{1}{\omega_\lambda(\boldsymbol{q})}\coth\frac{\hbar\omega_\lambda(\boldsymbol{q})}{2k_\mathrm{B}T}\,.\end{aligned} \tag{12.3.5}$$

In terms of the density of states this reads

$$\left\langle \boldsymbol{u}^2(m)\right\rangle = \frac{\hbar V}{2NM}\int \mathrm{d}\omega\,\frac{g(\omega)}{\omega}\coth\frac{\hbar\omega}{2k_\mathrm{B}T}\,. \tag{12.3.6}$$

At finite temperatures, for small values of ω, $\coth(\hbar\omega/2k_\mathrm{B}T)\approx 2k_\mathrm{B}T/\hbar\omega$, so when the small-$\omega$ asymptotic form (12.2.20) of the density of states is used, the integrand shows an ω^{d-3} dependence. Thus the above integral diverges at the lower limit even in $d=2$ dimensions. This indicates that while stable two-dimensional crystal structures may exist at $T=0$, they are destroyed at arbitrarily low nonzero temperatures.

This result can be considered as a consequence of the generalization of the *Mermin–Wagner theorem*,[5] originally formulated for magnetic systems. According to the general theorem, in two dimensions no long-range ordered state may exist at any finite temperature that breaks a continuous symmetry of the Hamiltonian. In our case the Hamiltonian is invariant under continuous translations, therefore in two- (or lower-) dimensional systems crystal structures that are invariant only under discrete translations cannot be stable at any finite temperature.[6] The melting point of the two-dimensional crystal is

[5] N. D. Mermin and H. Wagner, 1966.

[6] As mentioned above, in the one-dimensional case the crystalline state, which breaks the continuous symmetry, cannot be stable even as a ground state.

therefore $T_\mathrm{m} = 0$. The finite-temperature instability of the symmetry-breaking state is due to the excessively high number of soft Goldstone modes – acoustic phonons in our case – that are present even at very low temperatures and that disrupt the crystalline order.

Let us now calculate the mean square displacement in a three-dimensional crystal at finite temperature. Using the density of states (12.2.19) determined for the Debye model (with $p = 1$ for the monatomic basis),

$$\left\langle \boldsymbol{u}^2(m) \right\rangle = \frac{9\,\hbar}{M} \int\limits_0^{\omega_\mathrm{D}} \frac{\omega}{\omega_\mathrm{D}^3} \left[\frac{1}{\mathrm{e}^{\beta\hbar\omega} - 1} + \tfrac{1}{2} \right] \mathrm{d}\omega \,. \tag{12.3.7}$$

Instead of the Debye frequency ω_D the *Debye temperature* Θ_D defined via

$$k_\mathrm{B}\Theta_\mathrm{D} = \hbar\omega_\mathrm{D} \tag{12.3.8}$$

is commonly used. Changing to the variable $t = \beta\hbar\omega$,

$$\begin{aligned} \left\langle \boldsymbol{u}^2(m) \right\rangle &= \frac{9\,\hbar^2}{k_\mathrm{B}\Theta_\mathrm{D} M} \left(\frac{T}{\Theta_\mathrm{D}} \right)^2 \int\limits_0^{\Theta_\mathrm{D}/T} t \left[\frac{1}{\mathrm{e}^t - 1} + \tfrac{1}{2} \right] \mathrm{d}t \\ &= \frac{9\,\hbar^2}{k_\mathrm{B}\Theta_\mathrm{D} M} \left[\frac{1}{4} + \left(\frac{T}{\Theta_\mathrm{D}} \right)^2 \int\limits_0^{\Theta_\mathrm{D}/T} \frac{t}{\mathrm{e}^t - 1} \mathrm{d}t \right]. \end{aligned} \tag{12.3.9}$$

When only the first-order temperature correction to the zero-point vibrations is taken into account at low temperatures ($T \ll \Theta_\mathrm{D}$),

$$\left\langle \boldsymbol{u}^2(m) \right\rangle = \frac{9\,\hbar^2}{4k_\mathrm{B}\Theta_\mathrm{D} M} \left[1 + \frac{2\pi^2}{3} \left(\frac{T}{\Theta_\mathrm{D}} \right)^2 + \dots \right]. \tag{12.3.10}$$

Above the Debye temperature the integral can be evaluated using the expansion valid for $t < 1$,

$$\frac{t}{\mathrm{e}^t - 1} = 1 - \frac{t}{2} + \frac{t^2}{12} + \dots \,. \tag{12.3.11}$$

This gives

$$\left\langle \boldsymbol{u}^2(m) \right\rangle = \frac{9\,\hbar^2}{k_\mathrm{B}\Theta_\mathrm{D} M} \frac{T}{\Theta_\mathrm{D}} \left[1 + \frac{1}{36} \left(\frac{\Theta_\mathrm{D}}{T} \right)^2 + \dots \right]. \tag{12.3.12}$$

It seems reasonable to assume that the crystal melts when the root-mean-square displacement becomes comparable to the lattice constant a – that is, at the melting point T_m

$$\sqrt{\frac{9\,\hbar^2 T_\mathrm{m}}{k_\mathrm{B}\Theta_\mathrm{D}^2 M}} \approx \alpha a \,, \tag{12.3.13}$$

where α should be less than, but on the order of unity. Using the volume v of the primitive cell instead of the lattice constant,

$$T_{\mathrm{m}} \approx \frac{\alpha^2 k_{\mathrm{B}}}{9\hbar^2} \Theta_{\mathrm{D}}^2 M v^{2/3} \,. \tag{12.3.14}$$

Experimental data are in good agreement with the above relation if $\alpha \approx 1/4$. This is the *Lindemann criterion* for melting,[7] even though it was formulated somewhat differently before the advent of the Debye model.

12.3.2 The Specific Heat of the Phonon Gas

We shall now determine the contribution of atomic vibrations to the specific heat of solids. The previous expression for the thermal energy, (12.1.12), is expected to be modified by zero-point vibrations as

$$E = \sum_{\boldsymbol{q},\lambda} \hbar\omega_\lambda(\boldsymbol{q}) \left[\frac{1}{\mathrm{e}^{\beta\hbar\omega_\lambda(\boldsymbol{q})} - 1} + \tfrac{1}{2} \right] = \tfrac{1}{2} \sum_{\boldsymbol{q},\lambda} \hbar\omega_\lambda(\boldsymbol{q}) \coth\left(\tfrac{1}{2}\beta\hbar\omega_\lambda(\boldsymbol{q})\right) . \tag{12.3.15}$$

To demonstrate this, we shall make use of a statistical mechanical formula for the internal energy of a system with discrete energy levels E_i:

$$E = \frac{\sum_i E_i \mathrm{e}^{-\beta E_i}}{\sum_i \mathrm{e}^{-\beta E_i}} = -\frac{\partial}{\partial\beta} \ln \sum_i \mathrm{e}^{-\beta E_i} \,. \tag{12.3.16}$$

It is a simple matter to evaluate the sum on the right hand side, the partition function

$$Z = \sum_i \mathrm{e}^{-\beta E_i} \,. \tag{12.3.17}$$

Writing the energy of the phonon system as a sum of individual modes,

$$E_i = \sum_{\boldsymbol{q},\lambda} E_i(\boldsymbol{q},\lambda) = \sum_{\boldsymbol{q},\lambda} \hbar\omega_\lambda(\boldsymbol{q})[n_i(\boldsymbol{q},\lambda) + \tfrac{1}{2}] \,, \tag{12.3.18}$$

where $n_i(\boldsymbol{q},\lambda)$ can be any nonnegative integer. Then

$$\begin{aligned} Z &= \prod_{\boldsymbol{q},\lambda} \sum_i \mathrm{e}^{-\beta E_i(\boldsymbol{q},\lambda)} \\ &= \prod_{\boldsymbol{q},\lambda} \left(\mathrm{e}^{-\beta\hbar\omega_\lambda(\boldsymbol{q})/2} + \mathrm{e}^{-3\beta\hbar\omega_\lambda(\boldsymbol{q})/2} + \mathrm{e}^{-5\beta\hbar\omega_\lambda(\boldsymbol{q})/2} + \dots \right) \\ &= \prod_{\boldsymbol{q},\lambda} \frac{\mathrm{e}^{-\beta\hbar\omega_\lambda(\boldsymbol{q})/2}}{1 - \mathrm{e}^{-\beta\hbar\omega_\lambda(\boldsymbol{q})}} \,. \end{aligned} \tag{12.3.19}$$

[7] F. A. Lindemann, 1910.

From this formula the above expression for the internal energy is indeed recovered. Alternatively, the internal energy of the phonon gas can be expressed in terms of its Helmholtz free energy

$$F = -k_{\mathrm{B}}T \ln Z = \sum_{\boldsymbol{q},\lambda} \frac{\hbar\omega_\lambda(\boldsymbol{q})}{2} + k_{\mathrm{B}}T \sum_{\boldsymbol{q},\lambda} \ln\left(1 - \mathrm{e}^{-\hbar\omega_\lambda(\boldsymbol{q})/k_{\mathrm{B}}T}\right)$$
$$= k_{\mathrm{B}}T \sum_{\boldsymbol{q},\lambda} \ln\left\{2\sinh\left(\frac{\hbar\omega_\lambda(\boldsymbol{q})}{2k_{\mathrm{B}}T}\right)\right\}, \qquad (12.3.20)$$

using the thermodynamic identity

$$E = F - T\frac{\partial F}{\partial T} = k_{\mathrm{B}}T^2 \frac{\partial \ln Z}{\partial T}, \qquad (12.3.21)$$

which is equivalent to (12.3.16). In terms of the density of states, the Helmholtz free energy and the internal energy read

$$F = k_{\mathrm{B}}TV \int \ln\left\{2\sinh\left(\tfrac{1}{2}\beta\hbar\omega\right)\right\}\, g(\omega)\,\mathrm{d}\omega\,, \qquad (12.3.22\text{-a})$$

$$E = \tfrac{1}{2}V \int \hbar\omega\, \coth(\tfrac{1}{2}\beta\hbar\omega)\, g(\omega)\,\mathrm{d}\omega\,. \qquad (12.3.22\text{-b})$$

Let us return to form (12.3.15) of the internal energy. At high temperatures the condition $\beta\hbar\omega_\lambda(\boldsymbol{q}) = \hbar\omega_\lambda(\boldsymbol{q})/k_{\mathrm{B}}T \ll 1$ holds for all frequencies of the spectrum, and so it is sufficient to keep the first two terms in the series expansion of $\coth x$ given in (3.2.83):

$$E = \sum_{\boldsymbol{q},\lambda} k_{\mathrm{B}}T\left[1 + \frac{1}{12}\left(\frac{\hbar\omega_\lambda(\boldsymbol{q})}{k_{\mathrm{B}}T}\right)^2 + \dots\right]. \qquad (12.3.23)$$

When the specific heat is calculated from this expression, the leading term gives the classical Dulong–Petit value. Its first correction would be of order $1/T^2$, however this is suppressed by other contributions that are neglected in the harmonic approximation.

The frequency of acoustic phonons usually ranges from zero to order $10^{12}\,\mathrm{Hz}$. As the phonon density of states is proportional to ω^2, acoustic phonons of frequency $10^{12}\,\mathrm{Hz}$ can be considered as typical. Even for optical phonons the frequency is at most $10^{13}\,\mathrm{Hz}$ – so typical phonon energies are between 10 and 100 meV. This also means that the Dulong–Petit law is expected to be valid only at temperatures well above 100 K. At lower temperatures energy and specific heat can be determined only numerically from the above expressions, and only when the spectrum is known. The energy can be split into a temperature-independent term

$$E_0 = \sum_{\boldsymbol{q},\lambda} \tfrac{1}{2}\hbar\omega_\lambda(\boldsymbol{q}) \qquad (12.3.24)$$

and a temperature-dependent part

$$E_T = \sum_{\boldsymbol{q},\lambda} \frac{\hbar\omega_\lambda(\boldsymbol{q})}{\mathrm{e}^{\beta\hbar\omega_\lambda(\boldsymbol{q})} - 1} \,. \tag{12.3.25}$$

Expressing the sum over the phonon quantum numbers by the density of states,

$$E_T = V \int \frac{\hbar\omega}{\mathrm{e}^{\beta\hbar\omega} - 1} g(\omega)\,\mathrm{d}\omega \,. \tag{12.3.26}$$

From the previous expressions for the energy, the specific heat is

$$\begin{aligned} C_V &= \sum_{\boldsymbol{q},\lambda} \frac{\partial}{\partial T} \frac{\hbar\omega_\lambda(\boldsymbol{q})}{\mathrm{e}^{\beta\hbar\omega_\lambda(\boldsymbol{q})} - 1} \\ &= k_\mathrm{B} \sum_{\boldsymbol{q},\lambda} \left(\frac{\hbar\omega_\lambda(\boldsymbol{q})}{2k_\mathrm{B}T} \right)^2 \sinh^{-2} \left(\frac{\hbar\omega_\lambda(\boldsymbol{q})}{2k_\mathrm{B}T} \right) , \end{aligned} \tag{12.3.27}$$

or alternatively

$$\begin{aligned} C_V &= V \int \frac{\partial}{\partial T} \frac{\hbar\omega}{\mathrm{e}^{\beta\hbar\omega} - 1}\, g(\omega)\,\mathrm{d}\omega \\ &= V k_\mathrm{B} \int \left(\frac{\hbar\omega}{2k_\mathrm{B}T} \right)^2 \sinh^{-2} \left(\frac{\hbar\omega}{2k_\mathrm{B}T} \right) g(\omega)\,\mathrm{d}\omega \,. \end{aligned} \tag{12.3.28}$$

At low temperatures, where practically only low-energy acoustic phonons (which belong to the linear regime of the dispersion curve) are excited, the Debye model provides a good approximation. Using the Debye model density of states in (12.3.26),

$$E_T = 9pN \int_0^{\omega_\mathrm{D}} \frac{\hbar\omega}{\mathrm{e}^{\beta\hbar\omega} - 1} \frac{\omega^2}{\omega_\mathrm{D}^3}\,\mathrm{d}\omega \,. \tag{12.3.29}$$

With the new variable $t = \beta\hbar\omega$,

$$E_T = 9pNk_\mathrm{B}T \left(\frac{k_\mathrm{B}T}{\hbar\omega_\mathrm{D}} \right)^3 \int_0^{\hbar\omega_\mathrm{D}/k_\mathrm{B}T} \frac{t^3}{\mathrm{e}^t - 1}\,\mathrm{d}t \,. \tag{12.3.30}$$

Besides the factor p, which specifies the number of atoms in the basis, all information about the material characteristics of the crystal is contained in the parameter ω_D. Using the Debye temperature Θ_D instead,

$$E_T = 9pNk_\mathrm{B}T \left(\frac{T}{\Theta_\mathrm{D}} \right)^3 \int_0^{\Theta_\mathrm{D}/T} \frac{t^3}{\mathrm{e}^t - 1}\,\mathrm{d}t \,. \tag{12.3.31}$$

Introducing the *Debye function* through the definition

$$D_3(x) = \frac{3}{x^3} \int_0^x \frac{t^3}{\mathrm{e}^t - 1} \, \mathrm{d}t \,, \tag{12.3.32}$$

we have

$$E_T = 3pNk_\mathrm{B}TD_3\left(\frac{\Theta_\mathrm{D}}{T}\right). \tag{12.3.33}$$

The specific heat is then

$$C_V = 3pNk_\mathrm{B}\left[D_3\left(\frac{\Theta_\mathrm{D}}{T}\right) - \frac{\Theta_\mathrm{D}}{T}D_3'\left(\frac{\Theta_\mathrm{D}}{T}\right)\right]. \tag{12.3.34}$$

A simple integration by parts shows that this is equivalent to the expression

$$C_V = 9pNk_\mathrm{B}\left(\frac{T}{\Theta_\mathrm{D}}\right)^3 \int_0^{\Theta_\mathrm{D}/T} \frac{t^4\mathrm{e}^t}{(\mathrm{e}^t - 1)^2} \mathrm{d}t \,. \tag{12.3.35}$$

When the condition $T \ll \Theta_\mathrm{D}$ is met, the energy and the specific heat contain the asymptotic form of the Debye function valid for large values of x. According to (C.2.14),

$$D_3(x) \approx \frac{3}{x^3}\frac{\pi^4}{15} \tag{12.3.36}$$

in this limit, and so

$$C_V = pN\frac{12\pi^4}{5}k_\mathrm{B}\left(\frac{T}{\Theta_\mathrm{D}}\right)^3. \tag{12.3.37}$$

In the low-temperature region the temperature dependence of the specific heat is usually well approximated by the cubic term. When C_V/T is plotted against T^2, as in Fig. 12.8, a straight line can be fit fairly well to the measured data.

While a pure T^3 temperature dependence is observed in ionic crystals that are electrical insulators, the finite intercept for metals indicates that their specific heat contains an additional term which is linear in T. We shall see in Chapter 16 (Volume 2) that this is due to the electrons that are not bound to the ion cores and move almost freely.

On the other hand, it follows from (C.2.13), the expansion of the Debye function valid for small values of x, that at high temperatures the Dulong–Petit value is recovered from (12.3.35). Since (12.3.34) and (12.3.35) give good approximations for the specific heat both in the low- and high-temperature regions, the numerical evaluation of the integral in the Debye function may give an interpolation formula that is valid in the intermediate region as well. As indicated in Fig. 12.9, experimental data for the temperature dependence of specific heat show remarkable agreement with this interpolation formula.

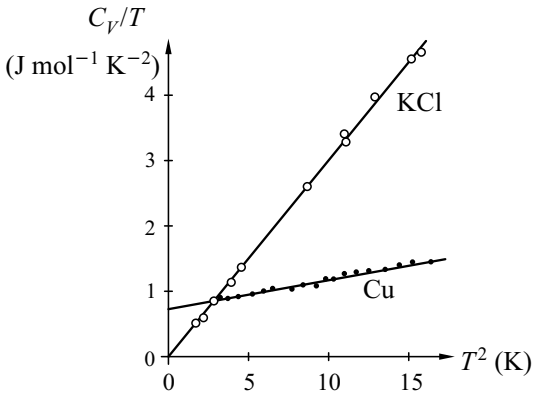


Fig. 12.8. Temperature dependence of specific heat for KCl and Cu at low temperatures, with C_V/T plotted against T^2 [based on P. H. Keesom and N. Pearlman, *Phys. Rev.* **91**, 1354 (1953), and H. M. Rosenberg, *Low Temperature Solid State Physics* (Oxford University Press, 1963)]

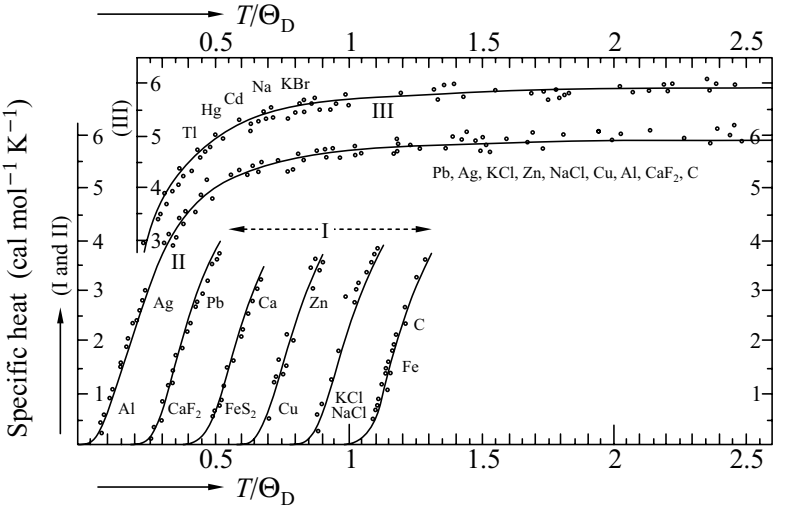


Fig. 12.9. Experimental data for the temperature dependence of the molar specific heat of some simple metals and ionic crystals, plotted against the reduced temperature T/Θ_D. For better visibility measured data are partially shifted, however the solid line is always the result of the Debye interpolation formula [M. Born and K. Huang, *Dynamical Theory of Crystal Lattices*, Oxford (1954)]

In principle, the Debye temperature Θ_D, which appears as a parameter in the specific heat, can be derived from the sound velocity. However, it is usually determined from specific heat data measured at low temperatures: once the electronic contribution is subtracted, (12.3.37) is fitted to the result.

The experimental values of the Debye temperature are listed for some elements and compounds in Table 12.1.

Table 12.1. Debye temperature (in kelvins) for some elements and compounds, determined by fitting (12.3.37) to experimental specific heat data

Element	Θ_D (K)	Element	Θ_D (K)	Compound	Θ_D (K)	Compound	Θ_D (K)
Na	158	Mn	410	LiF	732	CaF_2	510
K	91	Fe	467	NaCl	321	MgO	946
Cu	343	Ni	450	KCl	235	Fe_2O_3	660
Ag	225	C(d)	2230	RbCl	165	FeS_2	637
Au	165	C(g)	420	CsCl	175	SiO_2	470
Al	428	Ge	370	ZnS	315	Nb_3Sn	228

If the assumptions of the Debye model were valid for real materials, the interpolation formula could be fitted to the measured data with a single parameter over the entire temperature range. However, the agreement is not always as good as in the cases shown in Fig. 12.9. Therefore when the Debye formula is fitted to the specific heat data over a relatively small temperature range around T, the obtained Θ_D tends to be slightly different from the value derived from the low-temperature fit. It is in this sense that one may speak about the temperature dependence of the Debye temperature.

12.3.3 The Equation of State of the Crystal

A more complete analysis of the thermodynamic behavior of the crystal is based on the equation of state, which relates the pressure p, the volume V, and the temperature T. Starting with the free energy, pressure is given by the partial derivative

$$p = -\left(\frac{\partial F}{\partial V}\right)_T . \tag{12.3.38}$$

The free energy of the vibrating crystal is given by (12.3.20), therefore in the Debye model we have

$$\begin{aligned} F &= E_0 + 9\,p\,N\frac{k_B T}{\omega_D^3}\int_0^{\omega_D} \ln\left(1 - \mathrm{e}^{-\hbar\omega/k_B T}\right)\omega^2 \mathrm{d}\omega \\ &= E_0 + 9\,p\,N\,k_B T\left(\frac{T}{\Theta_D}\right)^3 \int_0^{\Theta_D/T} \ln\left(1-\mathrm{e}^{-t}\right) t^2\,\mathrm{d}t\,, \end{aligned} \tag{12.3.39}$$

where E_0 is the ground-state energy due to zero-point vibrations.

Volume dependence may occur through E_0 and Θ_D. Therefore the pressure is

$$p = -\frac{\partial E_0}{\partial V} + 3\,p\,N\,k_\mathrm{B}TD_3\left(\frac{\Theta_\mathrm{D}}{T}\right)\frac{1}{\Theta_\mathrm{D}}\frac{\partial\Theta_\mathrm{D}}{\partial V}\,. \tag{12.3.40}$$

Using (12.3.33) for the thermal energy E_T due to lattice vibrations,

$$p = -\frac{\partial E_0}{\partial V} + \gamma\frac{E_T}{V}\,, \tag{12.3.41}$$

where we have introduced the *Grüneisen parameter*[8]

$$\gamma = -\frac{V}{\Theta_\mathrm{D}}\frac{\partial\Theta_\mathrm{D}}{\partial V} = -\frac{\partial\ln\Theta_\mathrm{D}}{\partial\ln V}\,. \tag{12.3.42}$$

In harmonic crystals phonon frequencies are independent of the volume, and therefore so is the Debye temperature. This is most easily demonstrated by noting that in crystals of orthorhombic or higher symmetry the lattice constant a_i always appears in the combination $q_i a_i$ in the $\boldsymbol{q}$ dependence of the phonon frequencies, where q_i is an integral multiple of $2\pi/N_i a_i$ – so phonon frequencies are independent of the lattice constant. For a formal proof of general validity we shall assume that the crystal is strictly harmonic, therefore the relation

$$U(\boldsymbol{R}_m) = U_0 + \tfrac{1}{2}\sum_{m,n,\alpha,\beta} u_\alpha(\boldsymbol{R}_m)\Phi_{\alpha\beta}(\boldsymbol{R}_m - \boldsymbol{R}_n)u_\beta(\boldsymbol{R}_n) \tag{12.3.43}$$

is exact. For simplicity, we have assumed that each primitive cell contains a single atom. If the lattice constant increases from a to $(1+\epsilon)a$, the equilibrium position changes from $\boldsymbol{R}_m$ to $\overline{\boldsymbol{R}}_m = (1+\epsilon)\boldsymbol{R}_m$. If the actual ionic positions are $\boldsymbol{R}_m + \boldsymbol{u}(\boldsymbol{R}_m)$ in the original system and $\overline{\boldsymbol{R}}_m + \overline{\boldsymbol{u}}(\overline{\boldsymbol{R}}_m)$ with respect to the new equilibrium positions, we have

$$\boldsymbol{u}(\boldsymbol{R}_m) = \epsilon\boldsymbol{R}_m + \overline{\boldsymbol{u}}(\overline{\boldsymbol{R}}_m)\,. \tag{12.3.44}$$

Substituting this into the expression for the potential, and making use of the property that if $\overline{\boldsymbol{R}}_m$ is the equilibrium position then all terms that are linear in the displacement $\overline{\boldsymbol{u}}$ relative to it must vanish,

$$\begin{aligned} U(\boldsymbol{R}_m) = U_0 &+ \tfrac{1}{2}\epsilon^2\sum_{m,n,\alpha,\beta} R_m^\alpha\Phi_{\alpha\beta}(\boldsymbol{R}_m - \boldsymbol{R}_n)R_n^\beta \\ &+ \tfrac{1}{2}\sum_{m,n,\alpha,\beta}\overline{u}_\alpha(\overline{\boldsymbol{R}}_m)\Phi_{\alpha\beta}(\boldsymbol{R}_m - \boldsymbol{R}_n)\overline{u}_\beta(\overline{\boldsymbol{R}}_n). \end{aligned} \tag{12.3.45}$$

Since the force constants are the same as in the old system, so are the vibrational frequencies – that is to say frequencies are independent of the equilibrium volume:

[8] E. Grüneisen, 1926.

$$\frac{\partial \omega_\lambda(\boldsymbol{q})}{\partial V} = 0\,. \tag{12.3.46}$$

This implies that in the harmonic approximation the Grüneisen parameter must vanish, and so the equation of state is reduced to

$$p = -\frac{\partial E_0}{\partial V}\,. \tag{12.3.47}$$

The pressure is then independent of temperature:

$$\left(\frac{\partial p}{\partial T}\right)_V = 0\,, \tag{12.3.48}$$

and therefore the specific heats at constant volume and constant pressure, and the adiabatic and isothermal compressibilities, related by the thermodynamic identities

$$c_p = c_V - \frac{T\,(\partial p/\partial T)_V^2}{V\,(\partial p/\partial V)_T}, \qquad \frac{c_p}{c_V} = \frac{(\partial p/\partial V)_S}{(\partial p/\partial V)_T} \tag{12.3.49}$$

are equal. Even more interesting is that the linear thermal expansion coefficient of the solid, defined as

$$\alpha = \frac{1}{l}\left(\frac{\partial l}{\partial T}\right)_p = \frac{1}{3V}\left(\frac{\partial V}{\partial T}\right)_p\,, \tag{12.3.50}$$

where l is a characteristic linear dimension, vanishes in the harmonic approximation. To demonstrate this, consider the thermodynamic identity

$$\left(\frac{\partial p}{\partial T}\right)_V \left(\frac{\partial T}{\partial V}\right)_p \left(\frac{\partial V}{\partial p}\right)_T = -1 \tag{12.3.51}$$

and the bulk modulus K, the inverse of the isothermal compressibility κ defined as

$$K = \frac{1}{\kappa} = -V\left(\frac{\partial p}{\partial V}\right)_T\,; \tag{12.3.52}$$

the coefficient of thermal expansion may then be written as

$$\alpha = \frac{1}{3K}\left(\frac{\partial p}{\partial T}\right)_V\,. \tag{12.3.53}$$

According to (12.3.48), pressure is independent of temperature in the harmonic approximation, therefore the coefficient of thermal expansion vanishes. To account for the observed finite thermal expansion, we have to go beyond the harmonic approximation.

12.4 Anharmonicity

The possibility of describing the vibrations of the lattice in terms of independent collective excitations (phonons) stems from the applicability of the harmonic approximation to the crystal potential. Keeping only the second-order terms in the displacement, an exact diagonalization of the Hamiltonian was possible using the phonon creation and annihilation operators. When higher-order corrections become important, the harmonic approximation breaks down. Phonons will no longer be well-defined elementary excitations, as they can decay or be scattered by one another. Such processes change all the physical quantities that depend on the thermally excited states of the crystal.

12.4.1 Higher-Order Expansion of the Potential

To fourth order, the expansion of the potential is

$$\begin{aligned} U(\{\boldsymbol{u}(m,\mu\}) = U_0 &+ \frac{1}{2} \sum_{\substack{m,\mu,\alpha \\ n,\nu,\beta}} \frac{\partial^2 U}{\partial u_\alpha(m,\mu)\partial u_\beta(n,\nu)}\, u_\alpha(m,\mu) u_\beta(n,\nu) \\ &+ \frac{1}{3!} \sum_{\substack{m,\mu,\alpha,n,\nu,\beta \\ p,\pi,\gamma}} \frac{\partial^3 U}{\partial u_\alpha(m,\mu)\partial u_\beta(n,\nu)\partial u_\gamma(p,\pi)} \\ &\qquad \times u_\alpha(m,\mu) u_\beta(n,\nu) u_\gamma(p,\pi) \qquad (12.4.1) \\ &+ \frac{1}{4!} \sum_{\substack{m,\mu,\alpha,n,\nu,\beta \\ p,\pi,\gamma,r,\rho,\delta}} \frac{\partial^4 U}{\partial u_\alpha(m,\mu)\partial u_\beta(n,\nu)\partial u_\gamma(p,\pi)\partial u_\delta(r,\rho)} \\ &\qquad \times u_\alpha(m,\mu) u_\beta(n,\nu) u_\gamma(p,\pi) u_\delta(r,\rho) + \dots \ . \end{aligned}$$

When going beyond the second order, one has to go to the fourth order at least, since the third-order potential does not have an absolute minimum, and so the crystal would become unstable to large displacements otherwise. The fourth-order term stabilizes the potential. This problem does not arise in perturbation theory, however third- and fourth-order corrections may be of the same order because of the restrictions imposed on the third-order terms.

Since the corrections due to anharmonicity are expected to be small, we shall assume that it is sufficient to take anharmonicity into account only in a relatively low order of perturbation theory. Therefore we shall pass to the quantum mechanical discussion by neglecting the anharmonic term in the first step of the usual procedure. This means that we shall determine the normal modes from the dynamical matrix made up of the harmonic terms of the potential energy, the normal coordinates from the normal mode expansion of the displacement, and the canonical momenta from the Hamiltonian (i.e., the sum of the harmonic terms of the kinetic and potential energies), and then

impose canonical commutation relations on these coordinates and momenta. We then introduce phonon creation and annihilation operators, express the displacement in terms of them – as in (12.1.39) –, and use the quantized, operator-like quantity $\boldsymbol{u}$ in the third- and fourth-order terms of the expansion of the potential.

For notational simplicity, we shall consider a crystal with a monatomic basis, and write $u_\alpha(m)$ as u_m^α. The Hamiltonian of the anharmonic crystal is then

$$\mathcal{H} = \sum_{\boldsymbol{q}\lambda} \hbar\omega_\lambda(\boldsymbol{q}) \left(a_\lambda^\dagger(\boldsymbol{q})a_\lambda(\boldsymbol{q}) + \tfrac{1}{2}\right) + \frac{1}{3!} \sum_{\substack{m,n,p \\ \alpha\beta\gamma}} \frac{\partial^3 U}{\partial u_m^\alpha \partial u_n^\beta \partial u_p^\gamma} u_m^\alpha u_n^\beta u_p^\gamma + \frac{1}{4!} \sum_{\substack{m,n,p,r \\ \alpha\beta\gamma\delta}} \frac{\partial^4 U}{\partial u_m^\alpha \partial u_n^\beta \partial u_p^\gamma \partial u_r^\delta} u_m^\alpha u_n^\beta u_p^\gamma u_r^\delta + \ldots , \tag{12.4.2}$$

where (12.1.39) should be used for the ionic displacements. The expansion coefficients are represented by double or triple Fourier series, exploiting the property that the potential depends only on the distance between lattice points. For example, for the third-order term

$$\Phi_3^{\alpha\beta\gamma}(m,n,p) = \frac{\partial^3 U}{\partial u_m^\alpha \partial u_n^\beta \partial u_p^\gamma} = \sum_{\boldsymbol{q}\boldsymbol{q}'} D_3^{\alpha\beta\gamma}(\boldsymbol{q},\boldsymbol{q}') \mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot(\boldsymbol{R}_m - \boldsymbol{R}_n)} \mathrm{e}^{\mathrm{i}\boldsymbol{q}'\cdot(\boldsymbol{R}_m - \boldsymbol{R}_p)} . \tag{12.4.3}$$

Since the displacement is the linear combination of the operators $a_\lambda(\boldsymbol{q})$ and $a_\lambda^\dagger(\boldsymbol{q})$, the Hamiltonian will contain eight terms that are cubic and sixteen that are quartic in them. To get a better idea of their structure, we write out in full detail a third-order term in which the annihilation operator is chosen in each $\boldsymbol{u}$:

$$\begin{aligned} &\frac{1}{3!} \sum_{\substack{m,n,p \\ \alpha\beta\gamma}} \sum_{\boldsymbol{q}\boldsymbol{q}'} D_3^{\alpha\beta\gamma}(\boldsymbol{q},\boldsymbol{q}') \mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot(\boldsymbol{R}_m - \boldsymbol{R}_n)} \mathrm{e}^{\mathrm{i}\boldsymbol{q}'\cdot(\boldsymbol{R}_m - \boldsymbol{R}_p)} \\ &\times \frac{1}{(2MN)^{3/2}} \sum_{\substack{\boldsymbol{q}_1\boldsymbol{q}_2\boldsymbol{q}_3 \\ \lambda_1\lambda_2\lambda_3}} \left(\frac{\hbar}{\omega_{\lambda_1}(\boldsymbol{q}_1)}\right)^{1/2} \left(\frac{\hbar}{\omega_{\lambda_2}(\boldsymbol{q}_2)}\right)^{1/2} \left(\frac{\hbar}{\omega_{\lambda_3}(\boldsymbol{q}_3)}\right)^{1/2} \\ &\times e_\alpha^{(\lambda_1)}(\boldsymbol{q}_1) e_\beta^{(\lambda_2)}(\boldsymbol{q}_2) e_\gamma^{(\lambda_3)}(\boldsymbol{q}_3) a_{\lambda_1}(\boldsymbol{q}_1) a_{\lambda_2}(\boldsymbol{q}_2) a_{\lambda_3}(\boldsymbol{q}_3) \mathrm{e}^{\mathrm{i}\boldsymbol{q}_1\cdot\boldsymbol{R}_m} \mathrm{e}^{\mathrm{i}\boldsymbol{q}_2\cdot\boldsymbol{R}_n} \mathrm{e}^{\mathrm{i}\boldsymbol{q}_3\cdot\boldsymbol{R}_p} . \end{aligned} \tag{12.4.4}$$

The sum over the lattice points vanishes unless

$$\boldsymbol{q}_1 + \boldsymbol{q} + \boldsymbol{q}' = \boldsymbol{G}_1 , \qquad \boldsymbol{q}_2 - \boldsymbol{q} = \boldsymbol{G}_2 , \qquad \boldsymbol{q}_3 - \boldsymbol{q}' = \boldsymbol{G}_3 , \tag{12.4.5}$$

where $\boldsymbol{G}_i$ is a reciprocal-lattice vector, that is

$$\boldsymbol{q}_1 + \boldsymbol{q}_2 + \boldsymbol{q}_3 = \boldsymbol{G} . \tag{12.4.6}$$

This term of the Hamiltonian corresponds to a process in which three phonons are absorbed simultaneously. The sum of the three wave vectors is then required to be equal to a reciprocal-lattice vector. Normal processes occur for $\boldsymbol{G} = 0$, while $\boldsymbol{G} \neq 0$ is associated with umklapp processes.

Terms that contain one creation and two annihilation operators, or two creation and one annihilation operator, or three creation operators also appear.

12.4.2 Interaction Among the Phonons

The possible processes that correspond to the various terms of the third-order expansion are illustrated diagrammatically in Fig. 12.10. Each wavy line represents a phonon. Moving from left to right, a line leaving a vertex corresponds to the emission of a phonon, while a line entering a vertex corresponds to the absorption of a phonon. The wave vectors cannot be chosen arbitrarily. As the consequence of discrete translational symmetry, the conservation of crystal momentum must hold in each process.

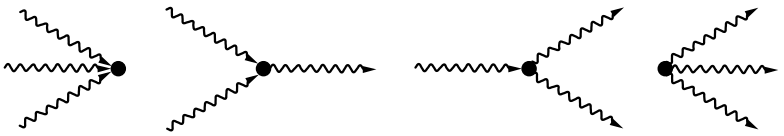


Fig. 12.10. Anharmonic processes with three phonons

The simultaneous absorption or emission of three phonons is important when calculating the total energy of the crystal but they are irrelevant to the physical properties that are of interest here. Such processes may exist because the intermediate (virtual) states live for a short time only, and, on account of the uncertainty principle, their energy may be arbitrary; energy conservation is required only when transition probabilities are calculated between an initial and a final state. Much more important are those terms in the Hamiltonian that contain one creation and two annihilation operators, or two creation and one annihilation operator.

These correspond to real physical processes in which two phonons merge into one or one phonon is split into two. When the conservation of crystal momentum and energy are imposed simultaneously, severe restrictions apply to the phonons that participate in the process. Suppose that in addition to

$$\boldsymbol{q}_1 + \boldsymbol{q}_2 = \boldsymbol{q}_3 + \boldsymbol{G} \tag{12.4.7}$$

the condition

$$\hbar\omega_{\lambda_1}(\boldsymbol{q}_1) + \hbar\omega_{\lambda_2}(\boldsymbol{q}_2) = \hbar\omega_{\lambda_3}(\boldsymbol{q}_3) \tag{12.4.8}$$

is also fulfilled. Using the Debye model it is immediately seen that only a triplet of collinear wave vectors can satisfy both conditions simultaneously.

For more realistic dispersion relations an even stricter restriction is obtained. As the dispersion curve of acoustic phonons is concave downward, energy and momentum cannot be simultaneously conserved in a process where the three phonons belong to the same acoustic branch.

Figure 12.11 shows the dispersion relation of acoustic phonons along a selected direction, assuming that transverse vibrations are degenerate. Plotting the dispersion curves once again, this time from the point that corresponds to the wave vector $\boldsymbol{q}_1$, the intersection points of the two families of curves give the possible values of $\boldsymbol{q}_2$ and $\boldsymbol{q}_3$. As longitudinal phonons are usually more energetic than transverse ones, a longitudinal phonon can easily decay into two transverse phonons or into a longitudinal and a transverse phonon, while the decay of transverse phonons may occur only via umklapp processes.

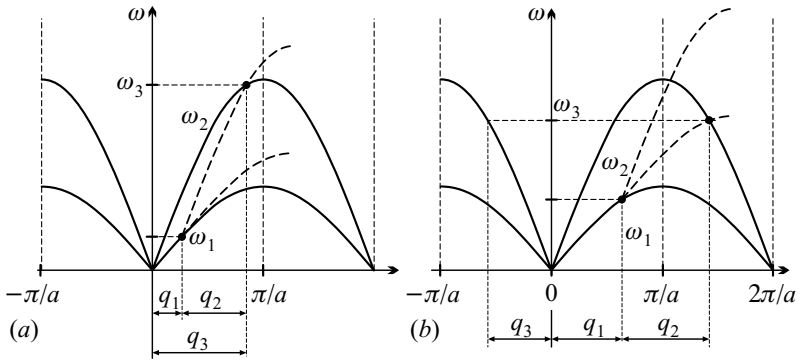


Fig. 12.11. Illustration of energy and momentum conservation for three-phonon processes: (*a*) normal process; (*b*) umklapp process

As phonons may decay as well as merge, they can no longer be considered as infinitely long-lived elementary excitations: their lifetime is necessarily finite. Because of the same processes, the energy of phonons is modified – renormalized – as a result of its interactions with other phonons. The contributions of such processes can be determined applying the methods used in many-body problems.

The role and contributions of fourth-order processes can be analyzed along the same lines. Simultaneous absorption and emission of four phonons are of little interest. The physically relevant processes are shown in Fig. 12.12.

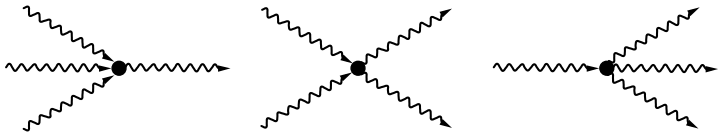


Fig. 12.12. Anharmonic processes with four phonons

12.4.3 Thermal Expansion and Thermal Conductivity in Crystals

In (12.3.53), using some general thermodynamic considerations, the coefficient of thermal expansion was written as

$$\alpha = \frac{1}{3K} \left(\frac{\partial p}{\partial T} \right)_V . \tag{12.4.9}$$

It was also demonstrated in the same section that α vanishes in the harmonic approximation. To examine the role of anharmonicity, we shall start with formula (12.3.41) that establishes the relationship between pressure and internal energy. This implies

$$\left(\frac{\partial p}{\partial T} \right)_V = \gamma \frac{C_V}{V} = \gamma \, c_V , \tag{12.4.10}$$

from which the *Grüneisen relation*[9] between the coefficient of thermal expansion and specific heat

$$\alpha = \gamma \frac{c_V}{3K} \tag{12.4.11}$$

is easily derived. Assuming as a first approximation a temperature-independent bulk modulus K, the coefficient of thermal expansion is found to be proportional to specific heat, that is, they exhibit the same temperature dependence.

Although the Grüneisen parameter γ vanishes in the harmonic approximation, in real crystals it does not because of the anharmonic terms. Substituting the experimental data for the coefficient of thermal expansion and specific heat into (12.4.11), a nonzero value is obtained for γ, which in general depends on the temperature. The Grüneisen relation nevertheless makes sense, since at low temperatures the coefficient of thermal expansion shows the same T^3 dependence as the specific heat, therefore the Grüneisen relation can be satisfied by a temperature-independent γ over a fairly broad temperature range. The situation is similar at room temperatures or above (i.e., at temperatures comparable to or higher than the Debye temperature), where the specific heat and the coefficient of thermal expansion are both constant. Table 12.2 shows the room-temperature Grüneisen parameter for a few elements.

When the thermal expansion coefficient is calculated for metals, the contribution of the electrons should not be ignored. We shall not revisit this question in Volume 2, where we focus on electrons, therefore it should be mentioned here that for the gas of free electrons

$$p_{\mathrm{el}} = \frac{2}{3} \frac{E_{\mathrm{el}}}{V} , \tag{12.4.12}$$

as we shall prove it in Chapter 16. This implies

$$\frac{\partial p_{\mathrm{el}}}{\partial T} = \frac{2}{3} \, c_V^{\mathrm{el}} . \tag{12.4.13}$$

[9] E. Grüneisen, 1926.

Table 12.2. The Grüneisen parameter for a few elements, measured at room temperature

Element	γ	Element	γ
Ag	2.40	Fe	1.60
Al	2.17	Co	1.87
Au	3.03	Si	0.45
Cu	1.96	Ge	0.73

The total thermal expansion coefficient is then

$$\alpha = \frac{1}{3K}\left[\gamma\, c_V^{\mathrm{ion}} + \frac{2}{3} c_V^{\mathrm{el}}\right]. \tag{12.4.14}$$

As we shall see, the specific heat of the electron gas is proportional to the temperature. At low temperatures, where the electronic contribution to the specific heat exceeds the phonon contribution, the coefficient of thermal expansion is proportional to T.

The previous considerations were based on the equation of state in the Debye model. For a more general interpretation of the Grüneisen parameter, the thermodynamic relations have to be applied to a system with a general dispersion relation. Starting from the free energy $F = E - TS$, the temperature dependence of pressure can be determined using (12.3.38). When the contribution of lattice vibrations is studied, (12.3.20) can be used for the free energy of independent phonons, with the assumption that phonon frequencies may be volume dependent:

$$p = -\frac{1}{2}\frac{\partial}{\partial V}\sum_{\boldsymbol{q},\lambda}\hbar\omega_\lambda(\boldsymbol{q}) + \sum_{\boldsymbol{q},\lambda}\left(-\frac{\partial}{\partial V}\hbar\omega_\lambda(\boldsymbol{q})\right)\frac{1}{\mathrm{e}^{\beta\hbar\omega_\lambda(\boldsymbol{q})}-1}. \tag{12.4.15}$$

The first term gives the contribution of the ground-state energy, and the second term accounts for the temperature dependence. Substituting this into (12.3.53) for the coefficient of thermal expansion,

$$\alpha = \frac{1}{3K}\sum_{\boldsymbol{q}\lambda}\left(-\frac{\partial}{\partial V}\hbar\omega_\lambda(\boldsymbol{q})\right)\frac{\partial}{\partial T}n_\lambda(\boldsymbol{q}). \tag{12.4.16}$$

To relate this formula to the form

$$c_V = \frac{1}{V}\sum_{\boldsymbol{q},\lambda}\hbar\omega_\lambda(\boldsymbol{q})\frac{\partial}{\partial T}n_\lambda(\boldsymbol{q}) = \sum_{\boldsymbol{q},\lambda} c_{V\lambda}(\boldsymbol{q}) \tag{12.4.17}$$

of the specific heat, we introduce a new quantity

$$\gamma_\lambda(\boldsymbol{q}) = -\frac{V}{\omega_\lambda(\boldsymbol{q})}\frac{\partial\omega_\lambda(\boldsymbol{q})}{\partial V} = -\frac{\partial\ln\omega_\lambda(\boldsymbol{q})}{\partial\ln V} \tag{12.4.18}$$

for each vibrational mode, which specifies the volume dependence of frequency for the mode in question. The Grüneisen relation (12.4.11) is satisfied if the parameter γ is defined as the weighted average of the $\gamma_\lambda(\boldsymbol{q})$:

$$\gamma = \frac{\sum_{\boldsymbol{q}\lambda} \gamma_\lambda(\boldsymbol{q}) c_{V\lambda}(\boldsymbol{q})}{\sum_{\boldsymbol{q}\lambda} c_{V\lambda}(\boldsymbol{q})} \,. \tag{12.4.19}$$

Collision (decay and merger) processes among phonons play an important role in rendering the thermal conductivity of nonmetallic solids finite. If the sample is not in thermal equilibrium but the temperature is different at the two ends, then a larger number of phonons are generated thermally at the higher-temperature end, and their motion toward the lower-temperature side may give rise to a heat current. This will be discussed in more detail in Chapter 24; only a simple picture is presented below.

We shall apply the results of the kinetic theory of gases to a gas of phonons. If the phonon mean free path Λ is finite because of the collisions among phonons, then the thermal conductivity is given by

$$\lambda_{\mathrm{ph}} = \tfrac{1}{3} \Lambda v c_V \,, \tag{12.4.20}$$

where v is the velocity of phonons, and c_V is the specific heat per unit volume. At high temperatures, where the Dulong–Petit law applies, the temperature dependence of thermal conductivity is governed by the temperature dependence of the phonon mean free path. Since the frequency of collision increases with the number of thermally generated phonons, it is plausible to assume that the mean free path is inversely proportional to the number of phonons. Above the Debye temperature the number of phonons

$$\langle n_\lambda(\boldsymbol{q}) \rangle = \frac{1}{\mathrm{e}^{\hbar\omega_\lambda(\boldsymbol{q})/k_{\mathrm{B}}T} - 1} \approx \frac{k_{\mathrm{B}}T}{\hbar\omega_\lambda(\boldsymbol{q})} \tag{12.4.21}$$

is proportional to T, so in this temperature range

$$\Lambda \propto 1/T \qquad \text{and} \qquad \lambda_{\mathrm{ph}} \propto 1/T \,. \tag{12.4.22}$$

At room temperatures the typical value of the mean free path ranges from 10^{-6} to 10^{-7} cm. At lower temperatures it can be substantially larger: around 20 K it is 10^{-2}–10^{-3} cm. At such low temperatures, where only long-wavelength (low-energy) phonons are generated thermally in significant numbers, one has to take into account that collisions do not reduce the heat current (energy current) of phonons as long as only normal scattering processes exist. This is because the dispersion relation of long-wavelength acoustic phonons is linear, and so momentum conservation goes hand in hand with energy conservation in the decays of such phonons. The dissipation of the heat current requires umklapp processes.

Further Reading

1. T. A. Bak, *Phonons and Phonon Interactions*, W. A. Benjamin, New York (1964).
2. M. Born and K. Huang, *Dynamical Theory of Crystal Lattices*, Oxford Classic Texts in the Physical Sciences, Oxford University Press, Oxford (1998).
3. P. Brüesch, *Phonons: Theory and Experiments*, Springer Series in Solid-State Sciences, Springer-Verlag, Berlin (1982).
4. A. A. Maradudin, E. W. Montroll, G. H. Weiss, I. P. Ipatova, *Theory of Lattice Dynamics in the Harmonic Approximation*, Second Edition, in: Solid State Physics, Supplement 3, Academic Press, New York (1971).
5. J. A. Reissland, *The Physics of Phonons*, John Wiley & Sons Ltd., London (1973).

13

The Experimental Study of Phonons

As we saw in the previous chapter, the excited states that correspond to the thermal motion of atoms in a crystalline solid can be interpreted in terms of independently generated collective elementary excitations (phonons) governed by the Bose–Einstein statistics. However, only rather indirect, integrated information on the phonon spectrum can be extracted from measurements of thermodynamic, macroscopic properties. In the present chapter we examine some direct methods for studying phonons.

13.1 General Considerations

In addition to giving an account of certain macroscopic properties of crystalline materials, the phonon gas model of lattice vibrations also provides a basis for the interpretation of the behavior of crystals under external influences. For example, when the solid is irradiated with electromagnetic radiation or a beam of particles, the interaction can be visualized as an interaction between the incident particles (quanta) and the gas of phonons. Although the processes that take place inside the system cannot be followed experimentally (the state of the system is not known precisely, therefore even the number of phonons before and after the measurement are unknown to the experimenter), and in many cases the details of the interactions with the phonons are also unknown, nevertheless experiments to measure the inelastic scattering or absorption of the incident radiation by the sample can provide information about the dynamics of the crystal lattice. The reason for this is simple. Consider radiation quanta or particles that are scattered or absorbed, as a result of their interaction with the system, and emit or absorb phonons in the process. By measuring the energy and momentum transferred to the scattered (absorbed) particle, and making use of the conservation of energy and crystal momentum, conclusions can be drawn about the phonon energy – that is, the dispersion relation –, as long as multiple scattering processes can be neglected. This may provide insight into the nature of the forces that hold the crystal together.

In principle any particle that interacts with the crystal in one way or another is suitable for such measurements. Without specifying whether photons, electrons, or neutrons are used, the equations for energy and momentum conservation can be written in a general form. We shall denote the particle energy before scattering by ε_i and the energy of the initial state of the phonon system by

$$E_\mathrm{i} = \sum_{\boldsymbol{q},\lambda} \hbar\omega_\lambda(\boldsymbol{q}) \left(n_\lambda(\boldsymbol{q}) + \tfrac{1}{2}\right). \tag{13.1.1}$$

In the scattering process the energy of the particle may change to ε_f, and the number of phonons with polarization λ and wave vector $\boldsymbol{q}$ to $n'_\lambda(\boldsymbol{q})$. Because of energy conservation,

$$\varepsilon_\mathrm{i} + \sum_{\boldsymbol{q},\lambda} \hbar\omega_\lambda(\boldsymbol{q}) n_\lambda(\boldsymbol{q}) = \varepsilon_\mathrm{f} + \sum_{\boldsymbol{q},\lambda} \hbar\omega_\lambda(\boldsymbol{q}) n'_\lambda(\boldsymbol{q}). \tag{13.1.2}$$

At the same time, on account of the discrete translational symmetry of the lattice, momentum conservatio is valid only up to a reciprocal-lattice vector:

$$\boldsymbol{p}_\mathrm{i} + \sum_{\boldsymbol{q},\lambda} \hbar\boldsymbol{q}\, n_\lambda(\boldsymbol{q}) = \boldsymbol{p}_\mathrm{f} + \sum_{\boldsymbol{q},\lambda} \hbar\boldsymbol{q}\, n'_\lambda(\boldsymbol{q}) + \hbar\boldsymbol{G}, \tag{13.1.3}$$

where $\boldsymbol{p}_\mathrm{i}$ and $\boldsymbol{p}_\mathrm{f}$ are the momenta of the incoming and scattered particles, and $\boldsymbol{G}$ is a reciprocal-lattice vector.

To be able to measure sufficiently precisely the energy and momentum transfer to the scattered particle, the energy and momentum of the scattered particle must be of the same order as the energy and quasimomentum of the phonons that are being studied. If one wishes to study the phonons over a relatively broad region of the Brillouin zone, the wavelength of the radiation must be of the order of the lattice constant. It was shown in Chapter 8 on diffraction that this part of the electromagnetic spectrum corresponds to X-ray photons.

Typical acoustic phonon energies can be identified with the thermal energy corresponding to the Debye temperature. According to the data listed in Table 12.1, the values of Θ_D are on the order of a few hundred kelvins, which corresponds to 10–100 meV. Phonon energies of the same order are obtained for optical phonons if the frequencies in Table 11.3 are converted to energies. This means that acoustic and optical phonon energies are typically a few tens of meV. X-ray photon energies are on the order of a keV, so phonons cause relatively little changes in their energy – therefore very high resolution devices are needed to extract the phonon dispersion relation from inelastic X-ray scattering. The energy of visible photons is around an electronvolt, therefore the issue of energy resolution is not so important. However, the wavelength being much larger than the lattice constant, practically only phonons around $\boldsymbol{q} = 0$ can be studied. On the other hand, for slow thermal neutrons both energy and momentum are in the suitable range. Before turning to the determination of the phonon dispersion relation through neutron scattering experiments, we examine what information can be extracted from optical measurements.

13.2 Optical Methods in the Study of Phonons

Although the methods discussed in this section had already been known in the first half of the 20th century, it was only with the advent of highly monochromatic laser sources that the scattering of light became a suitable and widely used method for studying lattice vibrations. This is because even tiny variations of the wavelength of the laser beam are easy to detect, and so sufficiently high resolutions can be achieved. A drawback of the method is that the wavelength of visible light ranges from $\lambda = 390$ to 750 nm, which corresponds to $q \sim 10^5\,\mathrm{cm}^{-1}$. This is several orders of magnitude smaller than the typical dimensions of Brillouin zones – therefore only phonons close to the center of the Brillouin zone ($q \approx 0$) can be studied with this method. The same applies, to an even greater degree, to using infrared radiation with wavelengths in the micron region.

In ionic crystals the electromagnetic field is coupled directly to the charged ions. The explicit form of the Hamiltonian will be given in Chapter 25. Here we shall only make use of the assertion that photons may create and annihilate phonons. In metals and semiconductors the indirect coupling of photons to phonons is more important. The electromagnetic field interacts directly with the highly polarizable and mobile electrons; then, in a second step, these electrons interact with the gas of phonons. We shall analyze these processes in more detail in Volume 2. Throughout the present chapter we shall refer to them as if they were direct interactions between the electromagnetic field and the phonons. We shall ignore the details of the interaction, and impose only the requirements of energy and momentum conservation in all real physical processes in which a photon is absorbed or scattered and a phonon is created or annihilated.

13.2.1 Infrared Absorption

In the simplest interaction between an electromagnetic field and a system described as a phonon gas, one photon of the incident beam is absorbed, and a phonon is created at the same time. As mentioned above, phonon energies are rarely in excess of a few tens of meV. As this is much smaller than the energy of visible photons (a few eV), such photons cannot be absorbed by the crystal via the emission of a single phonon. Because of energy conservation vibrating lattices can absorb only photons in the far infrared, where energies are much lower than in the visible range.

The frequency of the most energetic phonons is of order $10^{12}\,\mathrm{Hz}$. The wavelength of infrared photons of the same frequency is much larger than the lattice constant of crystals, i.e., their wave numbers are negligibly small compared to the dimensions of the Brillouin zone. Such photons can be absorbed only by long-wavelength phonons, with $\boldsymbol{q}$ being practically in the middle of the Brillouin zone, $q \approx 0$. Since long-wavelength acoustic phonons do not generate any electric dipole moment, only optical phonons participate in direct

interactions with the radiation field – more precisely, only transverse optical (TO) phonons, since the electric field $\boldsymbol{E}$ is perpendicular to the propagation direction, and so there are no interactions with longitudinal optical phonons. Since optical phonons of wave number $\boldsymbol{q} = 0$ can be regarded as the internal vibrations of a polyatomic basis, only certain transverse optical phonons may be created in infrared absorption processes – those that correspond to vibrations in which the electric dipole moment is changed. Phonons arising from such processes are called infrared active modes. As mentioned in Chapter 6, the transition matrix elements are subject to selection rules stipulated by the symmetries of the initial and final states. In our case the symmetries of the internal vibrational modes of the basis determine whether or not infrared absorption can occur in a particular mode. Even though the basis contains sixty atoms in fullerite (see page 30), there are only four infrared active optical phonons. There are thus four sharp absorption peaks in the infrared region, as shown in Fig. 13.1.

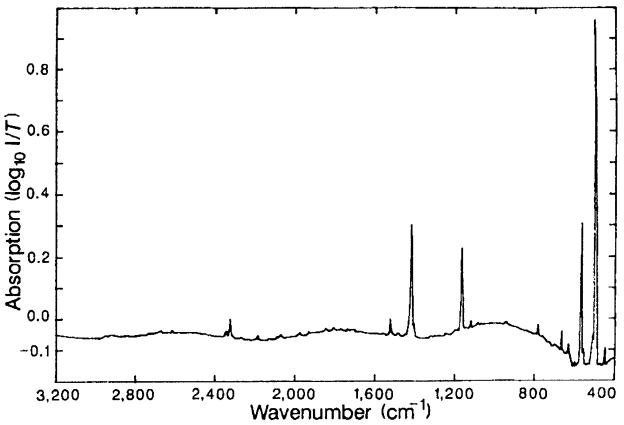


Fig. 13.1. Infrared absorption spectrum of an approximately 2 μm thick layer of solid C_{60} on a silicon substrate, referenced to a clean silicon substrate [W. Krätschmer et al., *Nature* **347**, 354 (1990)]

If only direct absorption processes existed, absorption would be observed only at the frequencies of TO phonons. However, due to phonon–phonon interactions, the TO phonon created in the photon absorption process may decay into two phonons or may be scattered by another phonon. These possibilities are shown in Fig. 13.2.

Since the intermediate virtual phonon does not need to obey the law of energy conservation, and only the initial and final states have to be of the same energy, absorption is observed in a finite range of energies. This is illustrated in Fig. 13.3, where the reflectivity of $LiGaO_2$ is plotted in the infrared region.

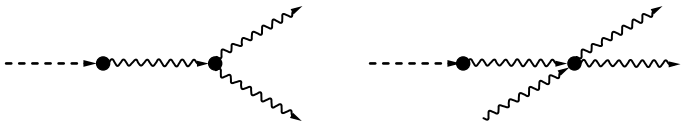


Fig. 13.2. Second-order processes, in which the TO phonon created in the photon absorption process splits into two phonons or is scattered by another phonon. Dashed lines represent photons, and wavy lines phonons

At wave numbers associated with certain well-defined frequencies reflectivity drops abruptly, indicating a strong absorption by the sample.

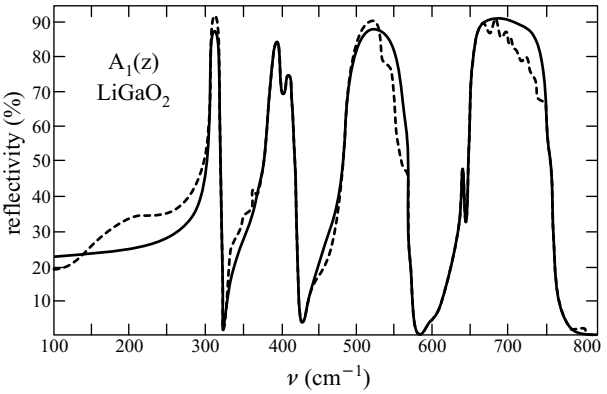


Fig. 13.3. Variations of the reflectivity of $LiGaO_2$ in the infrared region. The dashed line represents measured data; the solid line is the result calculated from the frequencies of longitudinal and transverse optical phonons [P. Knoll and H. Kuzmany, *Phys. Rev. B* **29**, 2221 (1984)]

13.2.2 Raman Scattering

In the lowest order of interaction, which is linear both in the electromagnetic field and in the ionic displacement, only two processes are possible: the transformation of a photon into a phonon and vice versa. It may occur that the generated phonon decays into two because of the anharmonic terms, and one of the new phonons is then transformed back into a photon. The net result of this process is the transformation of the initial photon into a photon plus a phonon. This can be interpreted as the inelastic scattering of the photon by the lattice, where the energy difference is carried away by the emitted phonon. Alternatively, the phonon that appears in the first step merges with a thermally excited phonon of the system into a new phonon state, which is then transformed back into a photon. This process can be interpreted as an

inelastic photon scattering accompanied by a phonon absorption. A distinction is made between the scattering processes in which an acoustic or optical phonon is involved. The first phenomenon is called *Brillouin scattering*, and the second *Raman scattering*.

In Raman scattering[1] the emission or absorption of an optical phonon of wave vector $\boldsymbol{q} \approx 0$ gives rise to a change in the energy of the incident photon:

$$\hbar\omega' = \hbar\omega \mp \hbar\omega_\lambda(\boldsymbol{q} \approx 0)\,. \tag{13.2.1}$$

These processes are shown in Fig. 13.4.

Fig. 13.4. Two types of Raman scattering. In Stokes processes a phonon is emitted; in anti-Stokes processes a phonon is absorbed

The change in the photon energy is a direct measure of the energy of the optical phonon. It follows from (13.2.1) that the energy of the scattered phonons may have two values for each phonon branch. The energy of the photon re-emitted in the process with phonon emission is lower than the energy of the incident photon; this is the *Stokes component*. The process with phonon absorption gives rise to the *anti-Stokes component* at energies in excess of the incident photon energy. The intensity is temperature dependent for both peaks – but not in the same way. The intensity of the anti-Stokes line is proportional to the number of absorbable phonons. In thermal equilibrium the number of phonons with energy $\hbar\omega_\lambda(\boldsymbol{q})$ is given by the Bose–Einstein function

$$g_0(\omega_\lambda(\boldsymbol{q})) = \frac{1}{\mathrm{e}^{\hbar\omega_\lambda(\boldsymbol{q})/k_\mathrm{B}T} - 1}\,. \tag{13.2.2}$$

As there are no absorbable phonons at $T = 0$, the intensity of the anti-Stokes line vanishes there. The Stokes line is stronger, its intensity is finite even at $T = 0$, as the phonon emission probability is proportional to $g_0(\omega_\lambda(\boldsymbol{q})) + 1$ on account of stimulated emission processes. This is clearly shown in Fig. 13.5, where the Raman spectrum of silicon is presented at three different temperatures.

Just like for infrared absorption, rigorous selection rules – specified by the symmetries of the crystal – determine whether or not a particular phonon

[1] C. V. Raman, 1928, and independently of him G. Landsberg and F. Mandelstam, 1928. Sir Chandrasekhara Venkata Raman (1888–1970) was awarded the Nobel Prize in 1930 "for his work on the scattering of light and for the discovery of the effect named after him".

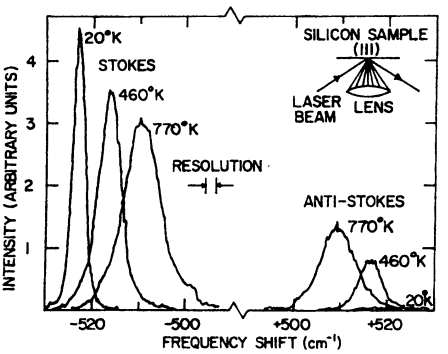


Fig. 13.5. Stokes and anti-Stokes components in the Raman spectrum of silicon at three different temperatures [T. R. Hart, R. L. Aggarwal and B. Lax, *Phys. Rev. B* **1**, 638 (1970)]

may be created or absorbed via Raman scattering. The general condition for Raman scattering is that the polarizability of the basis must be changed by the optical mode. Phonons that meet this condition are called *Raman active modes*. Consider, for example, fullerite. Due to the degeneracies that inevitably occur because of the high symmetry, the number of different vibrational frequencies at $\boldsymbol{q} = 0$ is 48 – however only 10 of them are Raman active. These are shown in Fig. 13.6.

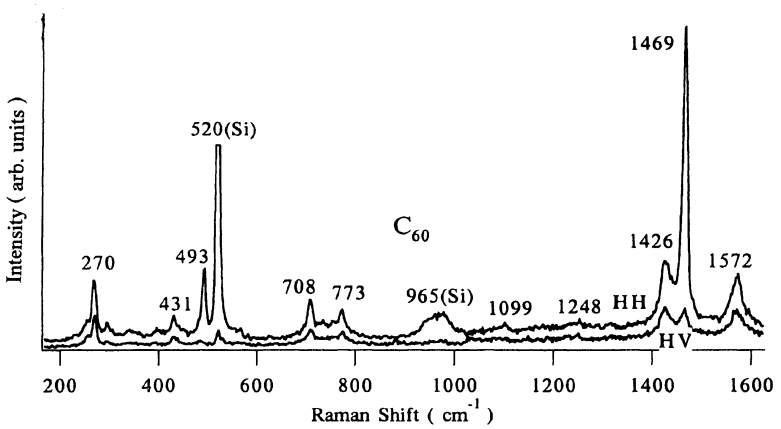


Fig. 13.6. Raman spectrum of a thin C_{60} layer. Two peaks due to the silicon substrate are observed in addition to the ten Raman active modes [P. Zhou et al., *Phys. Rev. B* **46**, 2595 (1992)]

The condition $\boldsymbol{q} \approx 0$ can be met not only via the creation or annihilation of a phonon with wave vector $\boldsymbol{q} \approx 0$ but also through scattering, in which a higher-order process gives rise to the creation or annihilation of a pair of phonons with opposite wave vectors ($\boldsymbol{q}$ and $-\boldsymbol{q}$). Since $\boldsymbol{q}$ is arbitrary, and the energy difference between the scattered and incident photons is now $2\hbar\omega_\lambda(\boldsymbol{q})$, a smeared-out continuum is obtained rather than a few sharp peaks. Nevertheless the largest contribution comes from the region where the density of states is high – that is, from optical phonons and high-energy acoustic phonons. This process is called the *two-phonon Raman scattering*. As it is a higher-order process, it could be expected to be less intense than one-phonon scattering, however, this can be compensated for by the high density of states of phonons at the boundary of the Brillouin zone. Figure 13.7 shows the calculated second-order Raman spectra for two alkali halide crystals.

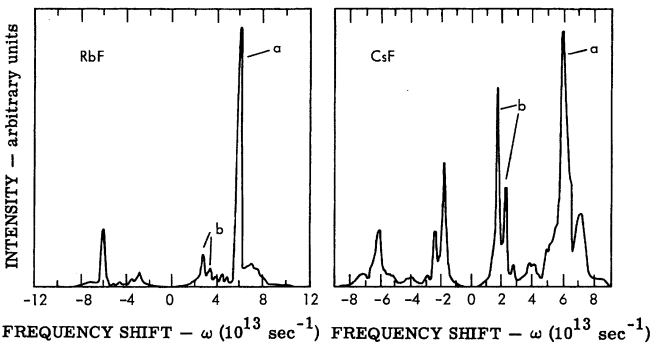


Fig. 13.7. Calculated second-order Raman spectra for two alkali halide crystals [J. R. Hardy and A. M. Karo, *Proc. of the Int. Conf. on Light Scattering Spectra of Solids*, New York, p. 99 (1968)]

13.2.3 Brillouin Scattering

The interpretation of *Brillouin scattering*[2] in terms of acoustic phonons is based on the conservation of energy (13.2.1) and crystal momentum; however, in contrast to Raman scattering, the change in the photon momentum cannot be neglected. It should be borne in mind that momentum conservation implies that the wave vector of the incident photon is changed – no matter how small the wave number of the emitted or absorbed phonon. Upon its penetration into a sample of refractive index n, the propagation velocity and thereby the wave number of the incident light is changed, the momentum associated with the photon is $\hbar n\boldsymbol{k}$. Momentum conservation then takes the form

[2] L. Brillouin, 1922.

$$\hbar n\boldsymbol{k}' = \hbar n\boldsymbol{k} \pm \hbar(\boldsymbol{q} + \boldsymbol{G})\,. \tag{13.2.3}$$

As it has been mentioned, the wave number $|\boldsymbol{k}|$ of photons is much smaller than the dimensions of the Brillouin zone, hence the previous equation can be satisfied only by the reciprocal-lattice vector $\boldsymbol{G} = 0$.

As the energies of visible photons and acoustic phonons are very different, on the order of 1 eV and 1–10 meV, respectively, there is hardly any change in the energy and the magnitude of the wave vector of the photon – however, the direction of the latter undergoes an observable change. We shall denote the angle between the propagation direction of the incident and scattered photons by θ; the vector triangle that illustrates momentum conservation is shown in Fig. 13.8.

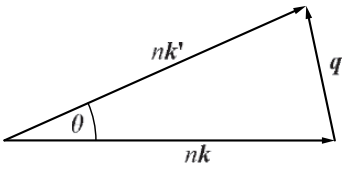


Fig. 13.8. The wave vector of the incident and scattered photons in Brillouin scattering

To a good approximation, the wave number of the emitted or absorbed phonon is

$$q = 2n\,k \sin\frac{\theta}{2}\,. \tag{13.2.4}$$

In terms of the photon frequency,

$$q = \frac{2n\,\omega}{c} \sin\frac{\theta}{2}\,. \tag{13.2.5}$$

In the long-wavelength region, which can be probed by Brillouin scattering, phonons of polarization λ propagate with the sound velocity c_λ. Since the phonon energy is the same as the change in the photon energy, we get

$$\Delta\omega = c_\lambda q = c_\lambda \frac{2n\,\omega}{c} \sin\frac{\theta}{2}\,. \tag{13.2.6}$$

When the intensity distribution of the scattered beam is measured as a function of energy in a fixed direction close to the direct beam, weaker peaks are observed on both sides of the large elastic peak, as shown in Fig. 13.9. They correspond to the emission and absorption of longitudinal and transverse phonons. The group velocities of the branches can be determined from the shift of the peaks relative to the direct beam.

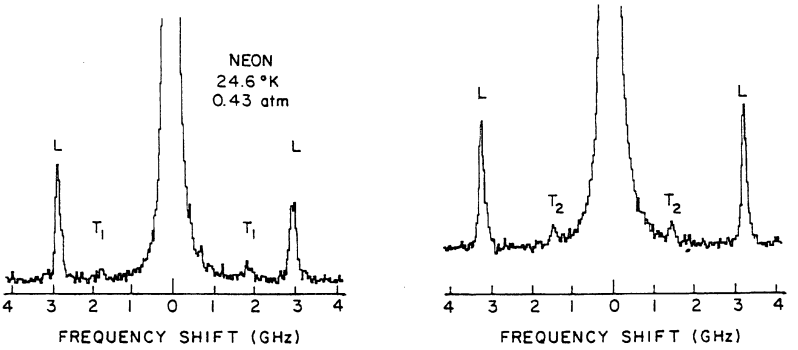


Fig. 13.9. Brillouin spectra of two single crystals of neon of different orientations, along a direction close to that of the direct beam [B. P. Stoichoff et al., *Proc. of the 2nd Int. Conf. on Light Scattering in Solids*, Paris, p. 450 (1971)]

13.3 Neutron Scattering on a Thermally Vibrating Crystal

In optical measurements only phonons at the center of the Brillouin zone ($\boldsymbol{q} \approx 0$) can be studied, since energy and momentum conservation can be satisfied only here on account of the great disparity between phonon and photon propagation velocities. The situation is different for neutrons. By rearranging (8.2.2), the formula relating the wave number and energy of the neutron we get

$$\varepsilon = 2.1 \left(k[\text{Å}^{-1}]\right)^2 \times 10^{-3}\,\text{eV}\,. \tag{13.3.1}$$

The energy of neutrons with a de Broglie wavelength of $\lambda \approx 1\,\text{Å}$ is about 0.1 eV. Expressed as a thermal energy in kelvins,

$$\varepsilon/k_\text{B} = 24 \left(k[\text{Å}^{-1}]\right)^2\,\text{K}\,. \tag{13.3.2}$$

Since the dimensions of a typical primitive cell are on the order of a few Å, the magnitude of the wave vector of the excitations (phonons) is typically on the order of a few Å^{-1}. The wave number of neutrons is in the same range if their thermal energy corresponds to about 100 K. Such so-called thermal neutrons are thus just in the ideal energy and wave-number range for inelastic scattering. Moreover, neutrons interact directly with nuclei, and – through magnetic interactions – with the electron cloud when the latter possesses a magnetic moment. Therefore thermal neutrons provide an almost ideal probe to study lattice vibrations and magnetic excitations in solids.[3]

[3] Bertram Neville Brockhouse (1918–2003) and Clifford Glenwood Shull (1915–2001) shared the Nobel prize in 1994 "for the development of neutron

Neutron emitted by reactors or spallation sources are usually many times more energetic than thermal neutrons. When they are passed through a moderator, a sufficiently intense beam of slow (thermal) neutrons is obtained. After collimation and monochromatization via Bragg reflection, a beam with definite energy and momentum emerges. If it is scattered inelastically on the sample, the task is to measure the angular and energy distribution of the scattered beam. Instead of the device used in neutron diffraction measurements (Fig. 8.9), a triple-axis spectrometer is now required. This is shown in Fig. 13.10.

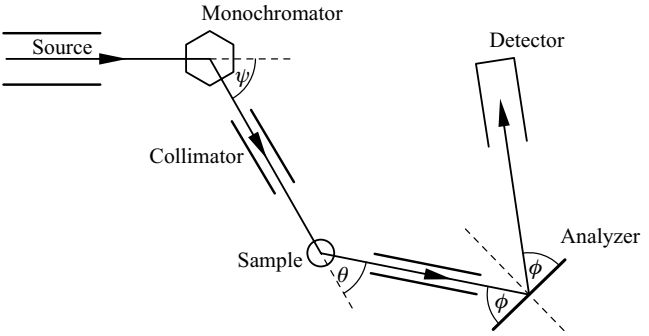


Fig. 13.10. Schematic sketch of a triple-axis spectrometer, used in inelastic neutron scattering measurements

The beam scattered by the sample falls on an analyzing crystal, which transmits only Bragg reflected neutrons of a particular wavelength (energy) to the detector. The monochromator, the sample holder, and the analyzing crystal can be rotated independently, which explains the origin of the name of the instrument.

13.3.1 Coherent Scattering Cross Section

The general discussion of neutron scattering on crystals with a polyatomic basis is straightforward in principle, however, for notational simplicity, we shall refrain from this, and assume that there is a single atom per primitive cell. We shall denote its position in the mth cell by $\boldsymbol{R}_m$.

The interaction between the neutron and the nucleus is described by the Fermi pseudopotential

$$U(\boldsymbol{r}-\boldsymbol{R}_m)=\frac{2\pi a_m\hbar^2}{m_{\mathrm{n}}}\delta(\boldsymbol{r}-\boldsymbol{R}_m)\,, \tag{13.3.3}$$

spectroscopy" (BNB) and "for the development of the neutron diffraction technique" (CGS).

where a_m is the scattering length of the atom located at $\boldsymbol{R}_m$. Even when the sample is chemically pure, the scattering length can be different for different isotopes of the same element. Based on the isotopic composition, an average scattering length may be introduced, and the sample can be conceived as if atoms with this average scattering length were located at each point of the lattice. This would lead to perfectly coherent scattering. However, in real samples an additional incoherent contribution appears in the scattering pattern because of the random deviations of the scattering length from the average value. The coherent and incoherent components can be fairly well separated experimentally. As we shall see, coherent scattering leads to sharp peaks in the intensity distribution of the scattered beam with respect to energy or scattering angle, while the incoherent part usually gives a featureless, smeared-out background. In our calculations we shall focus exclusively on the coherent contribution.

As demonstrated in the previous chapters on the structure of crystalline and amorphous materials, the intensity of the scattered beam is proportional to the $\boldsymbol{K} = \boldsymbol{k} - \boldsymbol{k}'$ component of the Fourier transform of the pair correlation function $g(\boldsymbol{r})$ that characterizes the spatial correlation of scattering centers. Apart from the contribution of the direct beam, this is just the structure factor

$$S(\boldsymbol{K}) = \frac{1}{N} \sum_{m,n} \left\langle \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{R}_m} \mathrm{e}^{\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{R}_n} \right\rangle . \tag{13.3.4}$$

For scattering on a thermally vibrating lattice it is plausible to assume that the time-dependent displacements $\boldsymbol{u}(m,t)$ need to be taken into account in the atomic position vectors, and therefore the cross section is related to the Fourier transform of the time-dependent correlation function, the *dynamical structure factor*

$$S(\boldsymbol{K},t) = \frac{1}{N} \sum_{m,n} \left\langle \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot(\boldsymbol{R}_m+\boldsymbol{u}(m,t))} \mathrm{e}^{\mathrm{i}\boldsymbol{K}\cdot(\boldsymbol{R}_n+\boldsymbol{u}(n)} \right\rangle . \tag{13.3.5}$$

It will be shown in Appendix E that this is indeed so. Here we just summarize the results derived there.

While neutrons are scattered elastically, without any energy transfer on a static system of atoms, scattering on a vibrating lattice is not necessarily elastic. According to the Van Hove formula, the doubly differential cross section – in which variations in the energy of the scattered particle are also taken into account – is proportional to the temporal Fourier transform of $S(\boldsymbol{K},t)$:

$$\begin{aligned} S(\boldsymbol{K},\omega) &= \frac{1}{N} \int_{-\infty}^{\infty} \mathrm{d}t\, \mathrm{e}^{\mathrm{i}\omega t} \sum_{m,n} \left\langle \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot(\boldsymbol{R}_m+\boldsymbol{u}(m,t))} \mathrm{e}^{\mathrm{i}\boldsymbol{K}\cdot(\boldsymbol{R}_n+\boldsymbol{u}(n))} \right\rangle \\ &= \frac{1}{N} \int_{-\infty}^{\infty} \mathrm{d}t\, \mathrm{e}^{\mathrm{i}\omega t} \sum_{m,n} \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot(\boldsymbol{R}_m-\boldsymbol{R}_n)} \left\langle \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{u}(m,t)} \mathrm{e}^{\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{u}(n)} \right\rangle , \end{aligned} \tag{13.3.6}$$

where ω and the energy transfer $\Delta\varepsilon$ are related by

$$\hbar\omega = \Delta\varepsilon = \frac{\hbar^2 \boldsymbol{k}^2}{2m_\mathrm{n}} - \frac{\hbar^2 \boldsymbol{k}'^2}{2m_\mathrm{n}} = E_\mathrm{f} - E_\mathrm{i}\,. \tag{13.3.7}$$

On account of energy conservation, this is the same as the change in the energy of the sample – that is, the energy of the created or annihilated phonon.

When atomic displacements are expressed, through (12.1.39), in terms of phonon creation and annihilation operators, the latter appear in the exponents of the structure factor. Because of their operator character, the exponentials cannot be merged; the relationship

$$\mathrm{e}^A \mathrm{e}^B = \mathrm{e}^{A+B} \tag{13.3.8}$$

holds only when A and B commute. When they do not, but their commutator $[A, B]$ commutes with A and B, then the *Baker–Hausdorff formula*[4] applies:

$$\mathrm{e}^A \mathrm{e}^B = \mathrm{e}^{A+B} \mathrm{e}^{\frac{1}{2}[A,B]}\,. \tag{13.3.9}$$

Since the thermodynamic average of such terms has to be taken, we shall make use of another theorem which asserts that for the linear combination

$$C = \sum_{\boldsymbol{q}} \left(\gamma_{\boldsymbol{q}} a(\boldsymbol{q}) + \gamma^*_{\boldsymbol{q}} a^\dagger(\boldsymbol{q})\right) \tag{13.3.10}$$

of the creation and annihilation operators of harmonic oscillators the thermal average of $\mathrm{e}^{\mathrm{i}C}$ is given by

$$\langle \mathrm{e}^{\mathrm{i}C} \rangle = \mathrm{e}^{-\frac{1}{2}\langle C^2 \rangle}\,. \tag{13.3.11}$$

When this is applied to $C = A + B$,

$$\langle \mathrm{e}^{\mathrm{i}(A+B)} \rangle = \mathrm{e}^{-\frac{1}{2}\langle A^2 + AB + BA + B^2 \rangle}\,. \tag{13.3.12}$$

From (13.3.9) we then have

$$\langle \mathrm{e}^{\mathrm{i}A} \mathrm{e}^{\mathrm{i}B} \rangle = \mathrm{e}^{-\frac{1}{2}\langle A^2 \rangle - \langle AB \rangle - \frac{1}{2}\langle B^2 \rangle} = \mathrm{e}^{-\frac{1}{2}\langle A^2 \rangle} \mathrm{e}^{-\langle AB \rangle} \mathrm{e}^{-\frac{1}{2}\langle B^2 \rangle}. \tag{13.3.13}$$

As suggested by expression (13.3.6) for $S(\boldsymbol{K}, \omega)$, we choose $A = -\boldsymbol{K} \cdot \boldsymbol{u}(m, t)$ and $B = \boldsymbol{K} \cdot \boldsymbol{u}(n)$. Since these quantities are the linear combinations of the phonon creation and annihilation operators with complex conjugate coefficients, and their commutator is a c-number, the above relations can be applied. For the first factor in (13.3.13) we introduce the notation

$$\mathrm{e}^{-W} = \mathrm{e}^{-\frac{1}{2}\langle (\boldsymbol{K} \cdot \boldsymbol{u}(m,t))^2 \rangle}. \tag{13.3.14}$$

In homogeneous crystals this quantity is independent of time and the equilibrium position $\boldsymbol{R}_m$ of the atom. Therefore the third factor gives the same

[4] H. F. Baker, 1905, and F. Hausdorff, 1906.

contribution. The combined factor of $\exp(-2W)$ is called the *Debye–Waller factor*.[5] This determines the intensity of the scattered beam. Its temperature dependence will be discussed in the next subsection.

The middle term in (13.3.13) is more interesting. As only those terms contribute to the thermal average in which the same phonon is created and annihilated, their contribution is

$$\exp\left\{\sum_{\boldsymbol{q},\lambda}\frac{\hbar}{2MN\omega_\lambda(\boldsymbol{q})}|\boldsymbol{K}\cdot\boldsymbol{e}^{(\lambda)}(\boldsymbol{q})|^2\Big[\langle n_\lambda(\boldsymbol{q})\rangle \mathrm{e}^{-\mathrm{i}[\boldsymbol{q}\cdot(\boldsymbol{R}_m-\boldsymbol{R}_n)-\omega_\lambda(\boldsymbol{q})t]} + (\langle n_\lambda(\boldsymbol{q})\rangle+1)\,\mathrm{e}^{\mathrm{i}[\boldsymbol{q}\cdot(\boldsymbol{R}_m-\boldsymbol{R}_n)-\omega_\lambda(\boldsymbol{q})t]}\Big]\right\}. \tag{13.3.15}$$

Collecting every term, we have

$$S(\boldsymbol{K},\omega)=\frac{1}{N}\int_{-\infty}^{\infty}\mathrm{d}t\,\mathrm{e}^{\mathrm{i}\omega t}\sum_{m,n}\mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot(\boldsymbol{R}_m-\boldsymbol{R}_n)}\mathrm{e}^{-2W} \times\exp\left\{\frac{\hbar}{2MN}\sum_{\boldsymbol{q},\lambda}\frac{|\boldsymbol{K}\cdot\boldsymbol{e}^{(\lambda)}(\boldsymbol{q})|^2}{\omega_\lambda(\boldsymbol{q})}\Big[\langle n_\lambda(\boldsymbol{q})\rangle\mathrm{e}^{-\mathrm{i}[\boldsymbol{q}\cdot(\boldsymbol{R}_m-\boldsymbol{R}_n)-\omega_\lambda(\boldsymbol{q})t]} + (\langle n_\lambda(\boldsymbol{q})\rangle+1)\,\mathrm{e}^{\mathrm{i}[\boldsymbol{q}\cdot(\boldsymbol{R}_m-\boldsymbol{R}_n)-\omega_\lambda(\boldsymbol{q})t]}\Big]\right\}. \tag{13.3.16}$$

The expression in the exponent of the last factor is usually small, therefore the exponential can be expanded in a power series. Regrouping the terms that depend on the coordinates of the ions,

$$S(\boldsymbol{K},\omega)=\frac{1}{N}\int_{-\infty}^{\infty}\mathrm{d}t\,\mathrm{e}^{\mathrm{i}\omega t}\mathrm{e}^{-2W}\left\{\sum_{m,n}\mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot(\boldsymbol{R}_m-\boldsymbol{R}_n)}\right. \tag{13.3.17}$$
$$+\frac{\hbar}{2MN}\sum_{mn}\sum_{\boldsymbol{q},\lambda}\frac{|\boldsymbol{K}\cdot\boldsymbol{e}^{(\lambda)}(\boldsymbol{q})|^2}{\omega_\lambda(\boldsymbol{q})}\Big[\langle n_\lambda(\boldsymbol{q})\rangle\mathrm{e}^{-\mathrm{i}(\boldsymbol{K}+\boldsymbol{q})\cdot(\boldsymbol{R}_m-\boldsymbol{R}_n)}\mathrm{e}^{\mathrm{i}\omega_\lambda(\boldsymbol{q})t}$$
$$\left.+\,(\langle n_\lambda(\boldsymbol{q})\rangle+1)\mathrm{e}^{-\mathrm{i}(\boldsymbol{K}-\boldsymbol{q})\cdot(\boldsymbol{R}_m-\boldsymbol{R}_n)}\mathrm{e}^{-\mathrm{i}\omega_\lambda(\boldsymbol{q})t}\Big]+\dots\right\}.$$

Performing the sum over the lattice points and the integral with respect to time,

$$S(\boldsymbol{K},\omega)=\mathrm{e}^{-2W}\sum_{\boldsymbol{G}}\Big\{\delta(\boldsymbol{K}-\boldsymbol{G})\delta(\omega) \tag{13.3.18}$$
$$+\frac{\hbar}{2MN}\sum_{\boldsymbol{q},\lambda}\frac{|\boldsymbol{K}\cdot\boldsymbol{e}^{(\lambda)}(\boldsymbol{q})|^2}{\omega_\lambda(\boldsymbol{q})}\Big[\langle n_\lambda(\boldsymbol{q})\rangle\delta(\omega+\omega_\lambda(\boldsymbol{q}))\delta(\boldsymbol{K}+\boldsymbol{q}-\boldsymbol{G})$$
$$+\,(\langle n(\boldsymbol{q})\rangle+1)\delta(\omega-\omega_\lambda(\boldsymbol{q}))\delta(\boldsymbol{K}-\boldsymbol{q}-\boldsymbol{G})+\dots\Big]\Big\}.$$

[5] P. Debye, 1914, and I. Waller, 1925.

Because of the presence of the factor $\delta(\omega)$ the first term corresponds to an elastic scattering with no energy transfer ($\Delta\varepsilon = 0$). Owing to the factor $\delta(\boldsymbol{K} - \boldsymbol{G})$, which arises from the sum over the lattice points, the scattered beam vanishes unless $\boldsymbol{K}$ is a vector of the reciprocal lattice. Thus the Bragg condition for diffraction is recovered. The other terms correspond to inelastic processes, in which phonons are emitted or absorbed.

13.3.2 Temperature Dependence of the Intensity of Bragg Peaks

In spite of thermal vibrations, Bragg peaks are not smeared out in the harmonic approximation. Their intensity is nonetheless reduced, and the Debye–Waller factor appears in their temperature dependence. Expressing once again atomic displacements, through (12.1.39), in terms of phonon creation and annihilation operators, and making use of the property that once again only those terms contribute to the thermal average in which the same phonon is created and annihilated,

$$\begin{aligned} 2W &= \left\langle (\boldsymbol{K}\cdot\boldsymbol{u}(m))^2 \right\rangle \\ &= \sum_{\boldsymbol{q},\lambda} \frac{\hbar}{2MN\omega_\lambda(\boldsymbol{q})} \left|\boldsymbol{K}\cdot\boldsymbol{e}^{(\lambda)}(\boldsymbol{q})\right|^2 \left\langle a_\lambda(\boldsymbol{q}) a_\lambda^\dagger(\boldsymbol{q}) + a_\lambda^\dagger(\boldsymbol{q}) a_\lambda(\boldsymbol{q}) \right\rangle \\ &= \sum_{\boldsymbol{q},\lambda} \frac{\hbar}{MN\omega_\lambda(\boldsymbol{q})} \left|\boldsymbol{K}\cdot\boldsymbol{e}^{(\lambda)}(\boldsymbol{q})\right|^2 \left(\left\langle n_\lambda(\boldsymbol{q})\right\rangle + \tfrac{1}{2}\right). \end{aligned} \tag{13.3.19}$$

In the isotropic case the branches that correspond to the three polarization directions are degenerate, and the polarization vectors of the longitudinal and two transverse branches can be taken as the basis vectors of a Cartesian coordinate system. Then summation over the polarization index λ of the branches leads to the expression

$$2W = \sum_{\boldsymbol{q}} \frac{\hbar K^2}{MN\omega(\boldsymbol{q})} \left(\left\langle n(\boldsymbol{q})\right\rangle + \tfrac{1}{2}\right). \tag{13.3.20}$$

Note that via (12.3.2) a very simple relation can be established between this quantity and the mean square displacement of ions:

$$2W = \frac{K^2}{3}\left\langle \boldsymbol{u}^2(m)\right\rangle. \tag{13.3.21}$$

When the phonon spectrum is taken in the Debye approximation, the mean square displacement – and with it, the Debye–Waller factor – can be evaluated. Using the formulas given in (12.3.10) and (12.3.12),

$$2W = \frac{3}{2}\frac{\hbar^2 K^2}{2M}\frac{1}{k_\mathrm{B}\Theta_\mathrm{D}}\left[1 + \frac{2\pi^2}{3}\left(\frac{T}{\Theta_\mathrm{D}}\right)^2 + \dots\right] \tag{13.3.22}$$

at low temperatures, and

$$2W = 6\frac{\hbar^2 K^2}{2M}\frac{1}{k_\mathrm{B}\Theta_\mathrm{D}}\left[\left(\frac{T}{\Theta_\mathrm{D}}\right) + \frac{1}{36}\left(\frac{\Theta_\mathrm{D}}{T}\right) + \dots\right] \tag{13.3.23}$$

at high temperatures. Since W appears in the exponent, the intensity of the Bragg peaks decreases with increasing temperature. The decrease in the scattering intensity occurs because neutrons emit virtual phonons and then (after an extremely short time) reabsorb them, or absorb phonons and after a short time re-emit them. Energy and momentum conservation is imposed only on the initial and final states but not on the intermediate virtual state – however, such processes modify the transition probability, and through it the scattered intensity. The probability of the process in which the phonon is first absorbed and then re-emitted is proportional to the thermal population of phonon states. The other process is proportional to $\langle n(\boldsymbol{q})\rangle + 1$ because of the possibility of stimulated emission. This explains the factor $2\langle n(\boldsymbol{q})\rangle + 1$ in W. Since an infinite succession of such processes can occur, W appears in the exponent in the Debye–Waller factor.

13.3.3 Inelastic Phonon Peaks

Having analyzed the first term in the expansion (13.3.18), which corresponds to elastic scattering (diffraction), let us now turn to the study of higher-order terms. These describe how additional peaks – which are due to processes with nonvanishing energy transfer – appear in the scattering pattern, at shifted positions relative to the Bragg peaks. Owing to the factor $\delta(\boldsymbol{K} + \boldsymbol{q} - \boldsymbol{G})$, the second term corresponds to a process in which the momentum of the scattered neutron is $\hbar\boldsymbol{k'} = \hbar\boldsymbol{k} + \hbar\boldsymbol{q} - \hbar\boldsymbol{G}$. The energy transfer in the scattering process can be written as $E_\mathrm{f} = E_\mathrm{i} - \hbar\omega_\lambda(\boldsymbol{q})$ because of the restriction $\Delta\varepsilon = E_\mathrm{f} - E_\mathrm{i} = -\hbar\omega_\lambda(\boldsymbol{q})$. This means that a phonon is absorbed in the scattering process. In Fig. 13.11 the phonon dispersion relation and $\varepsilon_\mathrm{n}(\boldsymbol{k} + \boldsymbol{q}) - \varepsilon_\mathrm{n}(\boldsymbol{k})$ are plotted against the component of $\boldsymbol{q}$ along the direction of $\boldsymbol{k}$. To account for the reciprocal-lattice vector that appears in the equation for crystal-momentum conservation, the phonon dispersion relation is repeated periodically, at equivalent wave vectors $\boldsymbol{q}$.

As the figure shows there are at least two intersection points, i.e., there are two $\boldsymbol{q}$s that satisfy both conditions simultaneously. Scattered neutrons must occur at the corresponding values of energy transfer. The intensity of this process is proportional to the thermal occupation $\langle n(\boldsymbol{q})\rangle$ of the phonon state that has to be annihilated. As (13.3.18) shows, the temperature-dependent Debye–Waller factor modifies the scattered intensity independently of the momentum and energy transfer – that is, the intensity of phonon peaks decreases with increasing temperature in the same way as the intensity of Bragg peaks.

The third term describes processes in which phonons are created. The presence of the factor $\delta(\boldsymbol{K} - \boldsymbol{q} - \boldsymbol{G})$ implies that the momentum of the neutron

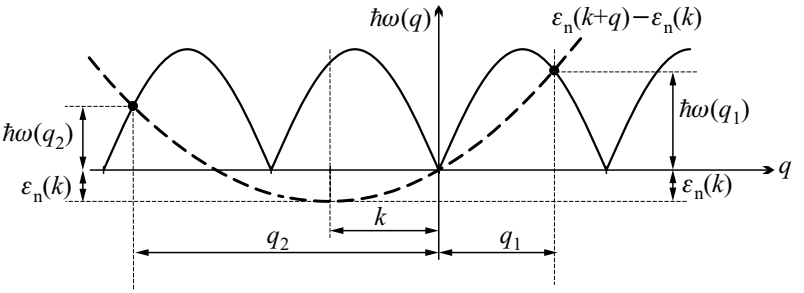


Fig. 13.11. Phonon and neutron dispersion relations for phonon absorption processes. The points of intersection indicate the solutions for which energy and quasi-momentum conservation are both satisfied

is changed to $\hbar\boldsymbol{k}' = \hbar\boldsymbol{k} - \hbar\boldsymbol{q} - \hbar\boldsymbol{G}$, while the restriction $\Delta\varepsilon = E_\mathrm{f} - E_\mathrm{i} = \hbar\omega_\lambda(\boldsymbol{q})$ shows that the energy of the scattering system is increased by the amount of the phonon energy: $E_\mathrm{f} = E_\mathrm{i} + \hbar\omega_\lambda(\boldsymbol{q})$. As this is a stimulated emission process, its intensity is proportional to $\langle n(\boldsymbol{q}) + 1\rangle$. The conditions for phonon emission processes are shown in Fig. 13.12. It can be demonstrated that emission occurs only above a certain threshold energy, which corresponds to the requirement that the neutron should possess more energy than the phonon that is to be created.

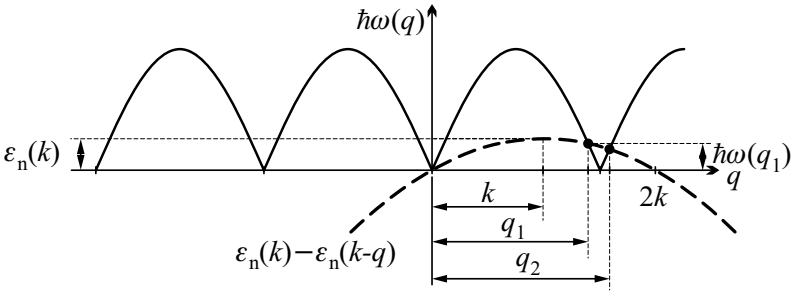


Fig. 13.12. Phonon and neutron dispersion relations for phonon emission processes. The points of intersection indicate the solutions for which energy and quasimomentum conservation are both satisfied

The equations for energy and crystal-momentum conservation impose severe restrictions on the energy and direction of the scattered beam. When neutrons of energy ε_n and wave vector $\boldsymbol{k}$ are incident on the sample, the above conditions are satisfied only at certain values of the energy for a given direction of the scattered beam. Using a triple-axis spectrometer, one can change the energy of the incident beam and its direction relative to the crystallographic axes of the sample. By measuring the energy of the scattered beam in

several directions, the changes in energy and crystal momentum permit the determination of the phonon spectrum.

In higher orders of the expansion (13.3.17) multiphonon processes are obtained with the corresponding restrictions of energy and momentum conservation. These usually give rise to a wide, smeared-out background.

13.3.4 The Finite Width of Phonon Peaks

In contrast to the infinitely narrow $\delta(\omega \mp \omega_\lambda(\boldsymbol{q}))$-type peaks of the previous description, peaks of finite width are observed in experiments, as shown in Fig. 13.13.

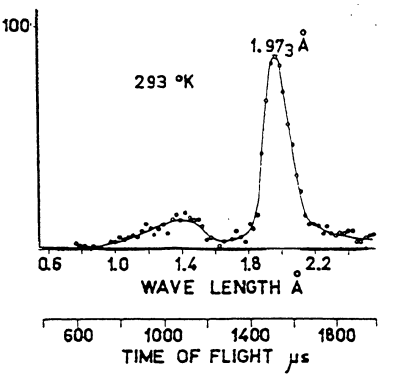


Fig. 13.13. The intensity of scattered neutrons as a function of energy for the one-phonon absorption peak in aluminum [K.-E. Larsson et al., *Proc. of the Symp. on Inelastic Scattering of Neutrons in Solids and Liquids*, Vienna, p. 587 (1960)]

The finite resolution of detectors is just one of the reasons. More important is that the phonon states obtained in the harmonic approximation are not exact eigenstates of the system. As we have seen, anharmonicity gives rise to phonon decay, leading to a finite phonon lifetime. Assume that the probability of finding the phonon decreases exponentially because of decay processes. Denoting the inverse lifetime by $\Gamma_\lambda(\boldsymbol{q})$, the time dependence of the phonon creation and annihilation operators are expected to be given by

$$\begin{aligned} a_\lambda(\boldsymbol{q},t) &= a_\lambda(\boldsymbol{q})\mathrm{e}^{-\mathrm{i}\omega_\lambda(\boldsymbol{q})t-\Gamma_\lambda(\boldsymbol{q})|t|}, \\ a_\lambda^\dagger(\boldsymbol{q},t) &= a_\lambda^\dagger(\boldsymbol{q})\mathrm{e}^{\mathrm{i}\omega_\lambda(\boldsymbol{q})t-\Gamma_\lambda(\boldsymbol{q})|t|} \end{aligned} \tag{13.3.24}$$

instead of (12.1.37). Therefore $S(\boldsymbol{K},\omega)$ in the cross section will have a form analogous to (13.3.17), however, an exponentially decreasing factor appears in the time integral because of the finite lifetime. Expression

$$\int_{-\infty}^{\infty} \mathrm{d}t\, \mathrm{e}^{\mathrm{i}\omega t} \mathrm{e}^{\pm \mathrm{i}\omega_\lambda(\boldsymbol{q})t - \Gamma_\lambda(\boldsymbol{q})|t|} = 2\frac{\Gamma_\lambda(\boldsymbol{q})}{(\omega \pm \omega_\lambda(\boldsymbol{q}))^2 + \Gamma_\lambda^2(\boldsymbol{q})} \tag{13.3.25}$$

shows that the inelastic peaks that correspond to phonon creation and annihilation are no longer sharp deltas: their intensity distribution is a Lorentzian curve. The smaller the parameter Γ (i.e, the longer the lifetime) the sharper the phonon peak. From the peak width the phonon lifetime can be determined quantitatively.

Further Reading

1. P. Brüesch, *Phonons: Theory and Experiments II, Experiments and Interpretation of Experimental Results*, Springer-Verlag, Berlin (1986).
2. W. Marshall and S. W. Lovesey, *Theory of Thermal Neutron Scattering: The Use of Neutrons for the Investigation of Condensed Matter*, Clarendon Press, Oxford (1971).
3. S. W. Lovesey, *Theory of Neutron Scattering from Condensed Matter*, Vol. I: *Nuclear Scattering*, Clarendon Press, Oxford (1984).

14

Magnetically Ordered Systems

A particularly interesting class of crystalline solids is constituted by those materials that exhibit regularity not only in the spatial arrangement of atoms but also in the alignment of their magnetic moments. It was shown in Chapter 3 that in a stand-alone atom or ion the orbital angular momentum $\boldsymbol{L}$ and the spin $\boldsymbol{S}$ of electrons on incomplete shells give rise to a magnetic moment $-g_{\mathrm{J}}\mu_{\mathrm{B}}\boldsymbol{J}$, where $\boldsymbol{J}$ is the dimensionless total angular momentum. Since the d- and f-electrons of ionic cores are usually localized in solids, too, these atomic moments are observed in the crystalline phase as well – although the crystalline field due to neighboring atoms may split the multiply degenerate ground state of the free ion, and therefore the moment may be modified. When the interaction between magnetic moments – namely, the quantum mechanical exchange interaction – is sufficiently strong on the scale of thermal energies, the magnetic moments of neighboring atoms may mutually align each other in some direction. It may occur that, in contrast to paramagnets, the orientation of each spin is rigidly fixed, however no correlation is observed in the orientation of magnetic moments, and the correlation function of spins drops off rapidly with distance. Since this static disorder is similar to the situation encountered in glasses, where atomic positions show a similar disorder, such systems are referred to as *spin glasses*. We shall analyze the properties of such materials in Chapter 36 of Volume 3. In the present chapter we shall be concerned with materials in which magnetic moments are aligned in some – usually crystallographically determined, high-symmetry – direction in such a way that their orientation shows long-range correlations. Such structures are said to be magnetically ordered.

First we shall get acquainted with the magnetic structure of ferromagnets, antiferromagnets, and ferrimagnets, which are the most common families of magnetically ordered materials. Then we shall examine the interactions that are responsible for magnetic ordering. At sufficiently high temperatures thermal fluctuations disrupt this order; the mechanism is similar to the melting of crystal lattices. First the simplest description provided by the mean-field theory will be presented, and then the behavior around the critical point will be

examined, using a phenomenological approach. After a brief discussion of the role of anisotropy in magnetic ordering we shall survey some properties of the domain structure, which plays an important role in technical applications of magnets. The quantum mechanical treatment of the behavior of magnetically ordered systems will be the subject of the next chapter.

14.1 Magnetic Materials

Atomic magnetic moments in paramagnetic materials become partially ordered upon the application of an external magnetic field. In their crystalline state the majority of the elements show this behavior. However, in certain 3d metals, rare-earth metals, and their compounds atomic magnetic moments become spontaneously ordered, aligned in a particular direction below a characteristic temperature. Based on the relative orientation of moments, ferromagnetic, antiferromagnetic, and ferrimagnetic materials are distinguished. Below we shall give a brief overview of these three types of magnetic materials, together with their most important properties.

14.1.1 Ferromagnetic Materials

A material is said to be ferromagnetic when the atomic magnetic moments of equal magnitude are aligned in a common direction by the interactions among them.[1] In the ground state this is perfectly true, however at higher temperatures alignment in a common direction is valid only in an average sense. Above a critical temperature T_C called the *Curie temperature* the order of the magnetic moments is completely destroyed.

Among transition metals with an incompletely filled 3d shell bcc iron (α-Fe) and fcc nickel become ferromagnetically ordered at relatively high temperatures, so they behave as magnetic materials at room temperature. Cobalt has two ferromagnetically ordered forms: ε-Co, which has an hcp structure and is stable at room temperature, and α-Co, which has an fcc structure and is stable at higher temperatures. Their magnetic transition temperatures (of the cubic structure for cobalt) and their saturation magnetization values are listed in Table 14.1. It is still common practice to express saturation magnetization in CGS units. To convert it to SI values, 1 gauss (CGS) is equivalent ($\widehat{=}$) to $10^3\,\mathrm{A/m}$ (SI). Comparison of different sources is sometimes difficult because some authors use the magnetic polarization $J = \mu_0 M$, while others prefer the saturation value of the magnetic moment per unit mass (σ_s).

In addition to the saturation magnetization, we have also given the moment μ per atom, which is calculated from it using Avogadro's number and the

[1] In a broader sense a system may be called ferromagnetic even when the ordered moments are not aligned completely parallel. The requirement is that one component should point in the same direction, while the perpendicular components must cancel out.

Table 14.1. The critical temperature T_C, saturation magnetization M_s, ordered moment μ, and two parameters that appear in the high-temperature susceptibility: the effective magneton number p_{eff}, and the paramagnetic Curie temperature Θ for the ferromagnetic 3d elements

Element	T_C (K)	M_s (10^3 A/m)	μ ($T = 0$) (μ_B)	p_{eff} ($T > T_c$)	Θ (K)
Fe	1043	1752	2.226	3.13	1101
Co	1388	1446	1.715	3.15	1415
Ni	627	5210	0.619	1.61	650

molar volume. The reader may realize that the values are different from those obtained for free ions. For comparison, we also give the magnetic moment – or more precisely, the effective magneton number p_{eff} obtained from the susceptibility measured in the high-temperature paramagnetic phase. Considering that the saturation moment is $|g_e|S\mu_B$ for localized spins of magnitude S, while the effective moment in the paramagnetic phase is $|g_e|(S(S+1))^{1/2}\,\mu_B$, the closeness of the low- and high-temperature values indicates that in metallic ferromagnets, too, the momentum that is responsible for magnetism comes predominantly from core electrons, rather than conduction electrons – the mobile electrons which are not bound to the ion cores, and which are responsible for metallic conduction. If this were not the case, the atomic magnetic moment would vanish at the same time as magnetization as the system is heated above the Curie temperature. Nevertheless the deviations from the free-ion values indicate that the ferromagnetic character of the elements of the iron group cannot be explained solely in terms of the localized moments of the ion cores: conduction electrons contribute as well. In this chapter we shall deal only with the model of localized spins. The magnetic properties due to mobile electrons will be discussed in Chapters 16 and 33.

Provided the moments are localized, their orientation in a magnetic structure can be represented by arrows drawn at the magnetic atoms. The ferromagnetic structures in simple, face-centered, and body-centered cubic crystals are shown in Fig. 14.1.

Since the moment of each magnetic atom points in the same direction in the ordered state, the translational symmetry of the lattice is not altered by the appearance of magnetic order. The symmetry of the magnetic structure is nonetheless lower than that of the nonmagnetic crystal because magnetic moments single out a preferred direction, and hence some of the rotational or reflection symmetries of the nonmagnetic state may be broken. For example, in cubic crystals magnetic moments may be aligned in the direction of a crystallographic axis. This axis is then no longer equivalent with the two others, and so the initial cubic symmetry is reduced to tetragonal.

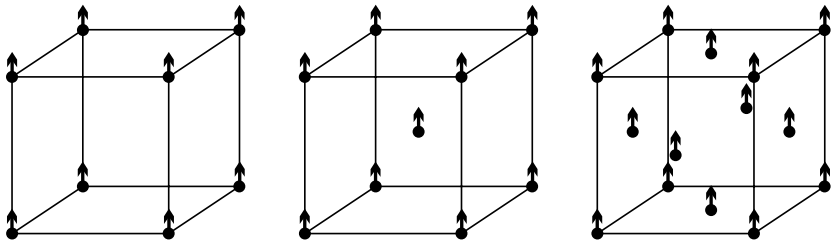


Fig. 14.1. Orientation of ordered magnetic moments in simple, face-centered, and body-centered cubic crystals

Ferromagnetic materials are found among the metallic and nonmetallic compounds of several elements with an incomplete 3d shell, not only the three listed in Table 14.1. Certain rare-earth metals and some of their compounds also show ferromagnetic properties at low temperatures. A handful of them are shown in Table 14.2, along with the critical temperature of the ferromagnetic phase and the saturation magnetization extrapolated to $T = 0$. As we shall see, the listed rare-earth metals are not purely ferromagnetic, as magnetic moments are not perfectly aligned in the ground state, but they possess a nonvanishing magnetization.

Table 14.2. Curie temperature and saturation magnetization (in units of 10^3 A/m) for some ferromagnetic rare-earth metals and their compounds, as well as ferromagnetic compounds containing 3d elements

Element	T_C (K)	M_s	Compound	T_C (K)	M_s	Compound	T_C (K)	M_s
Gd	293	2000	EuO	77	1910	Au_2MnAl	200	323
Tb	219	1440	EuS	16.5	1184	Cu_2MnAl	630	726
Ho	20	2550	$GdCl_3$	2.2	550	MnBi	620	675
Er	18		FeB	598		MnAs	670	870
Dy	85		Fe_2B	1043		$CrBr_3$	37	270

As it was mentioned on page 178, in 3d metals the orbital angular momentum is quenched, and so the localized moment comes entirely from spins. Consequently the electron g-factor, g_e appears in the moment. In rare-earth ions, where no such quenching occurs, the magnetic moment is calculated from the total angular momentum J and the Landé factor g_J. Measured data for the saturation magnetic moment are usually in excellent agreement with the expected value of the ordered moment for trivalent ions, calculated as $g_J \mu_B J$ from the data given in Table 3.6.

In each of the above compounds only one component has a nonvanishing magnetic moment. There exists a special group of ferromagnetic materials in

which two different types of atom possess a nonzero ordered moment. A few examples are given in Table 14.3.

Table 14.3. Curie temperature and ordered moment (in Bohr magnetons) for some ferromagnetic 3d compounds

Compound	T_C (K)	μ	μ
Fe_3Pt	400	$\mu_{Fe} = 2.7\,\mu_B$	$\mu_{Pt} = 0.5\,\mu_B$
CoPt	750	$\mu_{Co} = 1.6\,\mu_B$	$\mu_{Pt} = 0.25\,\mu_B$
$CoPt_3$	290	$\mu_{Co} = 1.64\,\mu_B$	$\mu_{Pt} = 0.26\,\mu_B$

14.1.2 Antiferromagnetic Materials

It was first suggested by L. Néel[2] in 1932 that in addition to ferromagnets with a macroscopic magnetization there may exist another class of materials, in which localized magnetic moments are aligned but the net spontaneous magnetization is nonetheless zero.

This happens because the magnetic moments of half of the atoms point in one direction, while those of the other half in the opposite direction. However, the cancellation of the magnetic moments occurs already on atomic scales, since the atoms with the two spin orientations are arranged in a checkerboard-like pattern; the magnetic moment of an atom is compensated by that of a nearest neighbor. Such materials are called *antiferromagnetic*. The transition temperatures T_N – called *Néel temperature* – of some antiferromagnetic materials are listed in Table 14.4. Comparison with Table 14.2 reveals that some rare-earth metals have both ferromagnetic and antiferromagnetic phases. In such cases the high-temperature paramagnetic phase is first transformed into an antiferromagnetic phase (the net spontaneous magnetization remains zero), and then, at a lower temperature, into the ferromagnetic phase (a nonzero net spontaneous magnetization appears).

The rare-earth metals listed in the first column crystallize in hexagonal close-packed structure; the oxides of composition AB in sodium chloride structure; the oxides and fluorides of composition ABC_3 in perovskite structure; the oxides of composition AB_2O_4 in spinel structure; the fluorides of composition XF_2 in tetragonal rutile structure – and antiferromagnetic materials of even more complex structure abound.

In contrast to ferromagnetic systems, many different antiferromagnetic structures are possible even in the simplest crystal lattices, as the requirement of zero net moment can be satisfied in various ways. We saw in Chapter 5 that

[2] Louis Eugène Félix Néel (1904–2000) was awarded the Nobel prize in 1970 "for [his] fundamental work and discoveries concerning antiferromagnetism and ferrimagnetism which have led to important applications in solid state physics".

Table 14.4. Néel temperature T_N for some antiferromagnetic materials

Element	T_N (K)	Compound	T_N (K)	Compound	T_N (K)
Cr	311	MnO	118	FeS	593
Mn	100	FeO	185	MnF_2	72
Ce	12.5	CoO	291	FeF_2	79
Nd	19.2	NiO	515	CoF_2	38
Sm	106	CuO	230	FeF_3	394
Eu	90.5	$NdFeO_3$	760	CoF_3	460
Dy	178	$LaFeO_3$	750	K_2NiF_4	97
Ho	132	$KMnF_3$	88	α-Fe_2O_3	948
Er	84	$KNiF_3$	275	Cr_2O_3	318
Tb	230	$NiCr_2O_4$	65	MnPt	975
Tm	56	$GeFe_2O_4$	10	Mn_3Pt	485

even when the restriction of collinearity is imposed – i.e., magnetic moments can only point in two opposite directions, up and down, for example –, a whole wealth of magnetic structures can still be found. Without going into details, we shall just present the antiferromagnetic structures that may occur in cubic crystals.

In one of the simplest antiferromagnetic structures occuring in simple cubic crystals with a monatomic basis the magnetic moments on adjacent lattice sites are directed oppositely. This structure, called type-G antiferromagnet, is shown in Fig. 14.2(c). The atoms with upward (downward) pointing spin form a face-centered cubic sublattice. The magnetic structure is therefore made up of two interpenetrating ferromagnetic fcc sublattices of opposite magnetization. The six nearest neighbors of each atom are on the other sublattice. Consequently the magnetic primitive cell is just the primitive cell of the fcc sublattice.

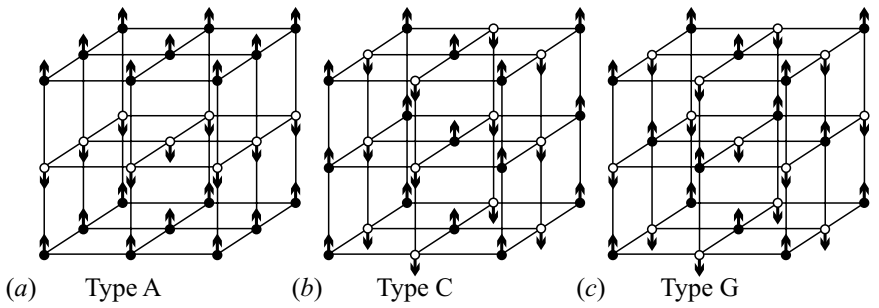


Fig. 14.2. Antiferromagnetic structures in simple cubic lattices

Other antiferromagnetic structures are also possible in a simple cubic crystal. Two of them are shown in parts (a) and (b) of Fig. 14.2. In the first case, called type-A antiferromagnet, magnetic moments are all parallel in the base plane but oppositely magnetized planes alternate along the perpendicular direction. The primitive translation vector of this magnetically ordered structure is twice as long in this direction as in the magnetically disordered state. In the second case, called type-C antiferromagnet, atoms with up and down spin are arranged in a checkerboard-like pattern in the base plane, and identical planes are stacked in the perpendicular direction. At first sight the primitive cell would seem to have been doubled in two directions. However, this is not so: the new primitive vectors are along the old face diagonals. The volume of the magnetic primitive cell is thus just twice the volume of the chemical primitive cell.

While in the previous case (part a) moments were aligned ferromagnetically in the planes (100), now (part b) they are aligned in the planes (110). In the perpendicular direction [110] oppositely magnetized planes alternate. Note that in the first case (part c) atomic magnetic moments are aligned ferromagnetically in the planes (111), and the magnetization direction alternates in subsequent planes. Therefore the magnetic moment density shows periodic variations along the direction [111].

Assuming that magnetic moments are strictly localized on atoms, its spatial distribution can be written as

$$\boldsymbol{\mu}(\boldsymbol{r}) = \sum_m \boldsymbol{\mu}(\boldsymbol{R}_m)\delta(\boldsymbol{r} - \boldsymbol{R}_m)\,, \tag{14.1.1}$$

where the sum is over the position of each atom with a magnetic moment. Function $\boldsymbol{\mu}(\boldsymbol{R}_m)$ is defined in discrete lattice points, therefore it can be expanded into a Fourier series according to (C.1.49):

$$\boldsymbol{\mu}(\boldsymbol{R}_m) = \frac{1}{\sqrt{N}} \sum_{\boldsymbol{q}} \boldsymbol{\mu}_{\boldsymbol{q}}\, \mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{R}_m}\,, \tag{14.1.2}$$

where the sum is over the wave vectors $\boldsymbol{q}$ inside the Brillouin zone. Note that the spatial distribution of the magnetic moment can be characterized by a single Fourier component $\boldsymbol{\mu}_{\boldsymbol{q}}$ for the antiferromagnetic structures presented above. In the three cases the vectors $\boldsymbol{q}$ are

$$\boldsymbol{q}_{\mathrm{A}} = (\pi/a)(0,0,1)\,, \quad \boldsymbol{q}_{\mathrm{C}} = (\pi/a)(1,1,0)\,, \quad \boldsymbol{q}_{\mathrm{G}} = (\pi/a)(1,1,1)\,. \tag{14.1.3}$$

These vectors of the reciprocal space correspond to high-symmetry points of the Brillouin zone: a face center, an edge center, and a vertex – that is, points X, M, and R in Fig. 7.2. Similar conclusions are drawn for the antiferromagnetic structures in body- and face-centered cubic lattices. This is not accidental – however its explanation requires the Landau theory of second-order phase transitions.

Up to now we have always spoken of up and down spins. The actual orientation of the moments with respect to the crystallographic axes – and thereby the possible directions of the vector $\boldsymbol{\mu}_{\boldsymbol{q}}$ – is determined by crystalline anisotropy. Energetically favorable directions usually correspond to some high-symmetry directions. This is why moments were drawn parallel to one of the crystallographic axes, although they can equally point in the direction of a face or body diagonal. We shall discuss this in more detail later.

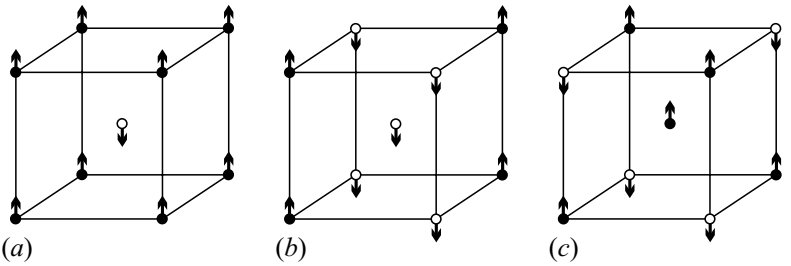


Fig. 14.3. Simple antiferromagnetic structures in body-centered cubic lattices

Three simple structures are possible in a body-centered cubic crystal; they are shown in Fig. 14.3. In structure (a) the atomic magnetic moments at the vertices point in the opposite direction as at the body centers. Atoms with upward and downward pointing moments form two interpenetrating simple cubic lattices. The nearest neighbors of each atom are on the opposite sublattice. Ferromagnetically coupled planes are perpendicular to the crystallographic axes. In structure (b) moments are coupled ferromagnetically along one edge of the cube, and antiferromagnetically along the two others. Ferromagnetic planes are now perpendicular to one of the face diagonals. In both magnetic structures the relative orientation of magnetic moments can be specified by a single vector $\boldsymbol{q}$, namely

$$\boldsymbol{q}_1 = (2\pi/a)(0,0,1), \quad \text{and} \quad \boldsymbol{q}_2 = (2\pi/a)(\tfrac{1}{2},\tfrac{1}{2},0)\,. \tag{14.1.4}$$

This is not the case for structure (c). Here the moment of the atom at a vertex of the cube is antiparallel to the moments of the six second-nearest neighbors, located at adjacent vertices. The relative orientation is characterized by the vector

$$\boldsymbol{q}_3 = (2\pi/a)(\tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2})\,. \tag{14.1.5}$$

This means that atomic moments are coupled ferromagnetically within planes (111), however, owing to the complex phase factor, the relative orientation of the moments at body centers with respect to those at vertices is not fixed. To understand the underlying physical reason consider the nearest neighbors of an atom: the moment points upward at four of them and downward at the four others. As the effects of the nearest neighbors cancel out, the direction of the

moment at the cube center is not specified by the moments at the vertices; instead, it is determined by the interactions with more distant neighbors. When the orientation of the moment at a particular cube center is fixed, then the moments at other cube centers cannot be chosen at will any more: an alternating pattern has to emerge. From a magnetic viewpoint the lattice can be decomposed into four face-centered cubic sublattices. The freedom in the choice of the orientation of the spin at the center manifests itself in the mathematical description as the requirement that both $\boldsymbol{q}_3$ and $-\boldsymbol{q}_3$ must appear in the Fourier representation of the moment density. The phases of the corresponding Fourier components determine the orientation of the atomic moment at the cell center. When the upper sign is chosen in

$$\begin{aligned}\boldsymbol{\mu}(\boldsymbol{R}_m) &= \frac{1}{\sqrt{2}}\boldsymbol{\mu}\left[\mathrm{e}^{\mathrm{i}(\boldsymbol{q}_3\cdot\boldsymbol{R}_m\pm\pi/4)} + \mathrm{e}^{-\mathrm{i}(\boldsymbol{q}_3\cdot\boldsymbol{R}_m\pm\pi/4)}\right] \\ &= \boldsymbol{\mu}\left[\cos(\boldsymbol{q}_3\cdot\boldsymbol{R}_m) \mp \sin(\boldsymbol{q}_3\cdot\boldsymbol{R}_m)\right] ,\end{aligned} \tag{14.1.6}$$

the structure shown in part (c) is obtained; when the lower sign is chosen, the spin at the center points in the opposite direction.

In terms of the primitive vectors (7.2.7) of the reciprocal of the bcc lattice the $\boldsymbol{q}$ are expressed as

$$\boldsymbol{q}_1 = \tfrac{1}{2}(\boldsymbol{b}_1 + \boldsymbol{b}_2 - \boldsymbol{b}_3) , \qquad \boldsymbol{q}_2 = \tfrac{1}{2}\boldsymbol{b}_3 , \qquad \boldsymbol{q}_3 = \tfrac{1}{4}(\boldsymbol{b}_1 + \boldsymbol{b}_2 + \boldsymbol{b}_3) . \tag{14.1.7}$$

It is readily seen in Fig. 7.7 that these vectors $\boldsymbol{q}$ are high-symmetry points – H, P, and N – at the boundary of the corresponding Brillouin zone.

The simple magnetic structures that occur in face-centered cubic crystals and their customary notation are shown in Fig. 14.4. For the overwhelming majority of antiferromagnetic materials with this type of Bravais lattice the alignment of the moments corresponds to the structures of the first or second type, although structures of the third and fourth types also exist in nature.

The spatial distribution of magnetic moments can be characterized by a single vector $\boldsymbol{q}$ in type-I and type-II antiferromagnets, while in type-III and type-IV structures both $\boldsymbol{q}_i$ and $-\boldsymbol{q}_i$ appear, in the form given by (14.1.6). Expressed in terms of the primitive vectors (7.2.12) of the reciprocal lattice, the vectors $\boldsymbol{q}$ characterizing the magnetic structure are

$$\begin{aligned}\boldsymbol{q}_{\mathrm{I}} &= (2\pi/a)(0,0,1) = \tfrac{1}{2}(\boldsymbol{b}_1 + \boldsymbol{b}_2) , \\ \boldsymbol{q}_{\mathrm{II}} &= (2\pi/a)(\tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2}) = \tfrac{1}{2}(\boldsymbol{b}_1 + \boldsymbol{b}_2 + \boldsymbol{b}_3) , \\ \boldsymbol{q}_{\mathrm{III}} &= (2\pi/a)(1,\tfrac{1}{2},0) = \tfrac{1}{4}(\boldsymbol{b}_1 + 2\boldsymbol{b}_2 + 3\boldsymbol{b}_3) , \\ \boldsymbol{q}_{\mathrm{IV}} &= (2\pi/a)(\tfrac{1}{2},\tfrac{1}{2},0) = \tfrac{1}{4}(\boldsymbol{b}_1 + \boldsymbol{b}_2 + 2\boldsymbol{b}_3) .\end{aligned} \tag{14.1.8}$$

By way of example, an even more complicated magnetic structure is also shown in Fig. 14.4. This could be called a four-q version of type-II antiferromagnets. It is derived from the simple type-II structure by flipping the atomic

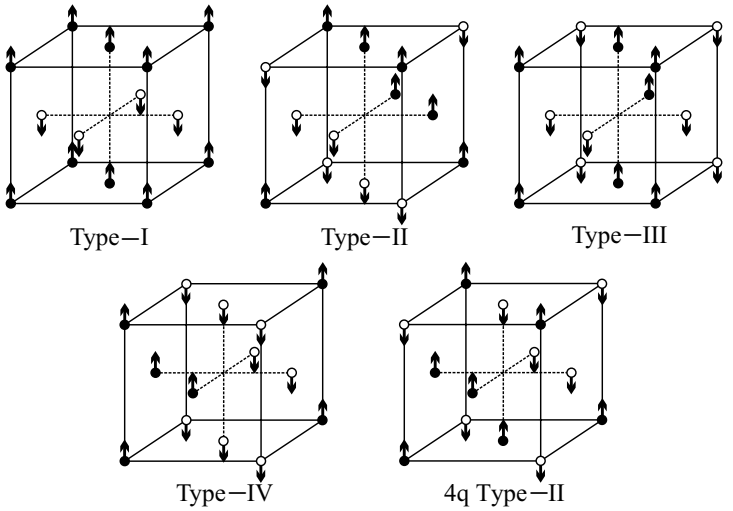


Fig. 14.4. Collinear magnetic structures in face-centered cubic crystals

moments at the face centers. This is possible because nearest-neighbor interactions do not determine the relative orientation of these moments with respect to the moments at the cube vertices. Specification of the structure requires four vectors

$$\begin{aligned} \boldsymbol{q} &= (2\pi/a)(\tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2})\,, & \boldsymbol{q}' &= (2\pi/a)(\tfrac{1}{2},\tfrac{1}{2},-\tfrac{1}{2})\,, \\ \boldsymbol{q}'' &= (2\pi/a)(\tfrac{1}{2},-\tfrac{1}{2},\tfrac{1}{2})\,, & \boldsymbol{q}''' &= (2\pi/a)(-\tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2})\,, \end{aligned} \tag{14.1.9}$$

and the magnetic moments can be given as

$$\boldsymbol{\mu}(\boldsymbol{R}_m) = \tfrac{1}{2}\boldsymbol{\mu}\left[-\mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{R}_m} + \mathrm{e}^{\mathrm{i}\boldsymbol{q}'\cdot\boldsymbol{R}_m} + \mathrm{e}^{\mathrm{i}\boldsymbol{q}''\cdot\boldsymbol{R}_m} + \mathrm{e}^{\mathrm{i}\boldsymbol{q}'''\cdot\boldsymbol{R}_m}\right]. \tag{14.1.10}$$

Apart from the type-IV structure, which occurs very rarely anyway, the antiferromagnetic structures are again represented by wave vectors $\boldsymbol{q}$ that correspond to high-symmetry points of the Brillouin zone, namely X, L, and W.

As it was mentioned in Chapter 8, neutron diffraction is ideally suited to the determination of magnetic structures. If neutrons interacted only with nuclei, the location of the diffraction peaks would permit only the determination of atomic positions, i.e., the crystal structure. However, through their spin and magnetic moment, neutrons also interact with electrons on incomplete shells that carry a nonzero moment. If these moments are arranged in an ordered magnetic structure, the Bragg condition of elastic magnetic scattering can only be met by the vectors of the reciprocal of the magnetic Bravais lattice, i.e., of the lattice spanned by the primitive translation vectors of the magnetic structure. Since the magnetic primitive cell may be larger than the chemical one, the primitive vectors of the magnetic reciprocal lattice may be shorter

than those determined by the nonmagnetic crystal structure. Therefore new Bragg peaks may appear at certain angles where no nuclear scattering occurs. Since they show different temperature dependence than the Bragg peaks arising from nuclear scattering, these *satellite peaks* permit the determination of the magnetic structure. An example is shown in Fig. 14.5.

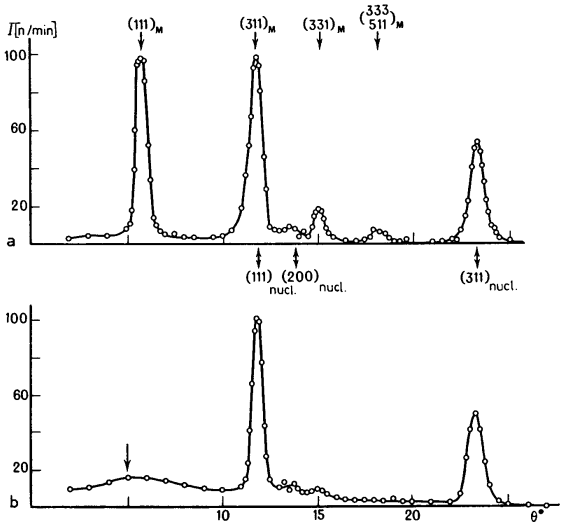


Fig. 14.5. Neutron-diffraction pattern of MnO: (*a*) in the low-temperature antiferromagnetically ordered phase at 80 K; (*b*) in the magnetically disordered phase at room temperature. The extra peaks that appear in the magnetic phase are satellites [C. G. Shull, W. A. Strauser, and E. O. Wollan, *Phys. Rev.* **83**, 333 (1951)]

When the diffraction peaks are indexed inside the chemical primitive cell, fractional numbers appear in the indices of the satellites. To ensure that only integers appear among the indices of the diffraction peaks, the magnetic primitive cell has to be used.

14.1.3 Spiral Magnetic Structures

In the foregoing we have examined magnetic structures in which adjacent magnetic moments are aligned either parallel or antiparallel. Such so-called collinear structures are by far the commonest, nevertheless there exist noncollinear structures as well in which the moments of adjacent atoms are at an angle to each other. Among such structures the most interesting are the spiral (helical or conical) structures observed in rare-earth metals. Four examples are shown in Fig. 14.6.

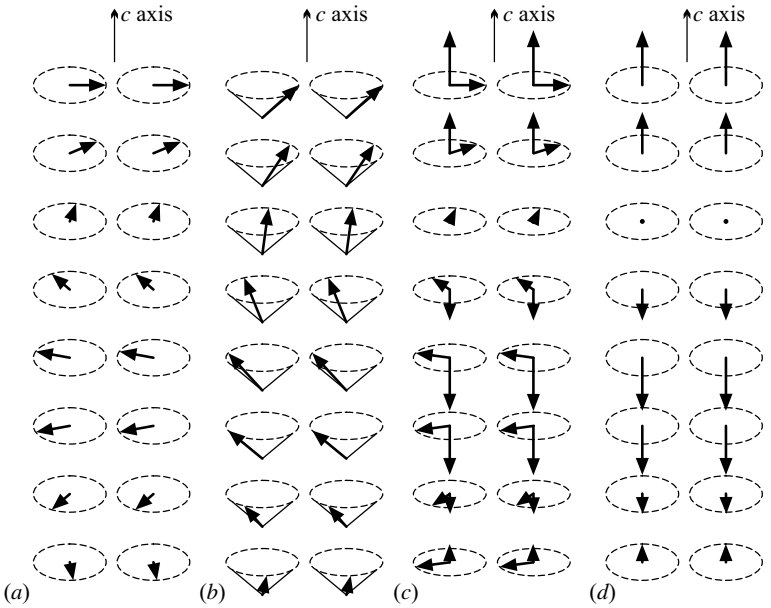


Fig. 14.6. Noncollinear magnetic structures in rare-earth metals: (*a*) the helical structure of dysprosium and holmium; (*b*) the conical ferromagnetic structure of holmium and erbium; (*c*) the sinusoidally modulated conical structure of erbium; (*d*) the longitudinal-wave structure of erbium

In the antiferromagnetic phase of dysprosium (between T_{N} and T_{C}) magnetic moments lie in the basal plane (the plane perpendicular to the sixfold axis) and are aligned ferromagnetically within each atomic plane perpendicular to the sixfold axis. However, the moments in adjacent layers make an angle φ with each other, and so the net magnetization of this helical structure is zero. The spatial variations of the magnetic moments can again be characterized by a single vector $\boldsymbol{q}$:

$$\mu_x(\boldsymbol{R}_m) = \mu_0 \cos(\boldsymbol{q} \cdot \boldsymbol{R}_m)\,, \quad \mu_y(\boldsymbol{R}_m) = \mu_0 \sin(\boldsymbol{q} \cdot \boldsymbol{R}_m)\,, \quad \mu_z(\boldsymbol{R}_m) = 0\,. \tag{14.1.11}$$

Here $\boldsymbol{q}$ is along the sixfold axis but it does not usually correspond to any high-symmetry point of the Brillouin zone. Consequently the periodicity of the magnetic structure can be incommensurate with the periodicity of the lattice. As we shall see in Chapter 29 of Volume 3, the underlying reason is that the oscillation period of the interaction governing magnetism in these systems is determined by the Fermi wave number. This is characteristic of the system of conduction electrons: in particular, it depends on the density of conduction electrons, and is of the same order as the linear dimensions of the Brillouin zone – but it may easily be incommensurate with it.

The same type of helical magnetic structure is observed in the antiferromagnetic phase of holmium below 132 K. Around 20 K another phase transition occurs, and a new (conical) magnetic structure appears in which the magnetic moment has a nonvanishing component along the sixfold axis. This component shows ferromagnetic order, whereas the perpendicular component is left practically unchanged. Thus the magnetic moment rotates on the surface of a cone whose axis is along the c-direction, as shown in Fig. 14.6(b).

Magnetic moments are not perfectly aligned in the low-temperature ferromagnetic phase of erbium, either. The magnetic moment has a nonvanishing ferromagnetic component along the sixfold axis, and another nonvanishing and rotating component perpendicular to it. Above the ferromagnetic phase two different antiferromagnetic phases are observed.

Between 18 and 52 K the component in the basal plane still rotates as one moves along the sixfold axis, but the axial component is also modulated sinusoidally with a wave vector that is the same as for the component in the basal plane. Finally, between 52 and 84 K the magnetic moments are oriented along the c-axis but their lengths oscillate sinusoidally. This pattern is called the longitudinal-wave structure.

14.1.4 Ferrimagnetic Materials

A magnetically ordered crystalline material is called ferrimagnetic if it is made up of two or more kinds of antiferromagnetically coupled ions whose magnetic moments are of different magnitude. Their existence was first pointed out by L. Néel in 1948. The magnetic moments are localized on rare-earth ions or trivalent iron ions, but similar magnetic materials can be obtained by arranging divalent and trivalent ions at the sites of two appropriately chosen sublattices in a complicated structure. Even when the magnetic primitive cell contains the same number of ions of each type – which is not always the case –, the two types of magnetic moment do not cancel perfectly in the antiparallel arrangement because of their different magnitudes. Table 14.5 shows the critical temperature for magnetic ordering for some ferrimagnetic materials.

Table 14.5. Magnetic ordering temperature for some ferrimagnetic materials

Ferrite spinels	T_C (K)	Garnets	T_C (K)
Fe_3O_4	860	$Y_3Fe_5O_{12}$	550
$CoFe_2O_4$	790	$Eu_3Fe_5O_{12}$	565
$NiFe_2O_4$	850	$Sm_3Fe_5O_{12}$	580
$MnFe_2O_4$	570	$Gd_3Fe_5O_{12}$	560
$CuFe_2O_4$	730	$(Y,Al)_3Fe_5O_{12}$	520
$(Ni,Al)Fe_2O_4$	2130	$(Y,Gd)_3Fe_5O_{12}$	520

A characteristic class of ferrimagnets is that of *ferrites* – compounds of composition $M^{2+}Fe_2^{3+}O_4^{2-}$ (where M stands for Co, Ni, Cu, Zn, or Cd) that crystallize in spinel or inverse spinel structure. A particular representative of this class is magnetite, Fe_3O_4, a compound that contains di- and trivalent iron ions and crystallizes in inverse spinel structure, as specified on page 224. The oxygen atoms form an fcc lattice. The tetrahedral sites of this lattice are occupied by Fe^{3+} ions; their moments are aligned. They are antiferromagnetically coupled to the moments of Fe^{2+} ions – which occupy the octahedral sites – and the remaining Fe^{3+} ions. The moment of divalent ions gives rise to the net magnetization below the critical temperature of magnetic ordering.

Another characteristic class of ferrimagnets is that of *rare-earth garnets*. Their general composition is $M_3^{3+}Fe_5^{3+}O_{12}^{2-}$, or in an alternative notation $3M_2O_3 \cdot 5Fe_2O_3$. These crystals belong to the space group O_h^{10} ($Ia\bar{3}d$), and their primitive cells contain eight $M_3Fe_5O_{12}$ units each, hence the Pearson symbol of this structure is cI160. Of the five Fe^{3+} ions per unit three have their moments aligned ferromagnetically; the two others are antiparallel to them, just like the moments of the rare-earth ions.

Because of the appearance of spontaneous magnetization, ferrimagnets could be expected to show a similar magnetic behavior to ferromagnets. However, the temperature dependence of the moment of rare-earth ions and iron ions may be essentially different in garnets. It may occur upon heating that the magnetization of the oppositely directed sublattices cancel out at the *compensation point* T_{comp}, and then a finite magnetization reappears at higher temperatures. Such an example is shown in Fig. 14.7.

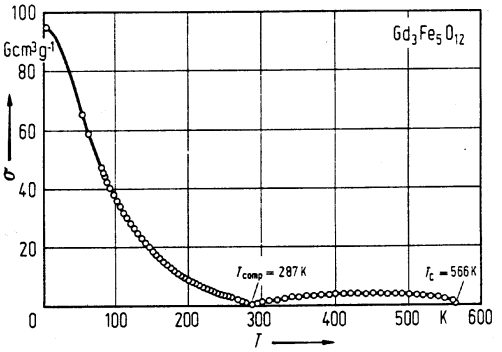


Fig. 14.7. The temperature dependence of the magnetization of gadolinium–iron garnet [E. E. Anderson, *Proc. of the Int. Conf. on Magnetism*, p. 660 (1964)]

14.2 Exchange Interactions

The dipole–dipole interaction

$$\mathcal{H}_{\text{dipole–dipole}} = \frac{\mu_0}{4\pi}\left[\frac{\boldsymbol{\mu}_1\cdot\boldsymbol{\mu}_2}{r^3} - \frac{3(\boldsymbol{\mu}_1\cdot\boldsymbol{r})(\boldsymbol{\mu}_2\cdot\boldsymbol{r})}{r^5}\right] \tag{14.2.1}$$

is always present between atomic magnetic moments. To evaluate the order of magnitude of the energy of this interaction, we shall assume that the magnetic moment is on the order of a Bohr magneton μ_{B}, while the distance r is on the order of the Bohr radius a_0. Making use of the formulas for μ_{B} and a_0 expressed in terms of more fundamental physical constants (see Appendix A), and the relation $\epsilon_0\mu_0 = 1/c^2$,

$$E \approx \frac{\mu_0}{4\pi}\frac{\mu_{\text{B}}^2}{a_0^3} = \left(\frac{e^2}{4\pi\epsilon_0\hbar c}\right)^2 \frac{e^2}{16\pi\epsilon_0 a_0}\,. \tag{14.2.2}$$

The first term is recognized as the square of the fine-structure constant $\alpha \approx 1/137$, and the second as the Rydberg energy, 13.6 eV. The energy of dipole–dipole interactions is therefore on the order of a meV at most. Such an energy cannot be responsible for the ordering of magnetic moments at temperatures as high as several hundred kelvins, as it could not stabilize any structure against the disruptive effects of thermal fluctuations. Dipole–dipole interactions play an important role only in magnetic anisotropy and the demagnetization of finite samples. Magnetism is due to another, much stronger interaction that is rooted in the Coulomb interaction but manifests itself only for quantum mechanical reasons: the exchange interaction.

14.2.1 Direct Exchange

As it was demonstrated in Section 4.4 on covalent bonds, even though the Coulomb interaction is spin independent, the singlet and triplet states of a two-electron hydrogen molecule are not of the same energy. According to (4.4.17), their energy difference is

$$\varepsilon_{\text{t}} - \varepsilon_{\text{s}} = 2C\left(N_-^2 - N_+^2\right) - 2I\left(N_-^2 + N_+^2\right) = 2\frac{C|S|^2 - I}{1-|S|^4}\,, \tag{14.2.3}$$

where S is the overlap integral between the wavefunctions of neighboring atoms, C is the two-particle interaction matrix element, and I is the exchange integral. These quantities are defined in (4.4.15) and (4.4.18).

When the overlap integral is small, the energy difference is determined by the exchange integral. It is therefore customary to call J, defined through the formula

$$\varepsilon_{\text{t}} - \varepsilon_{\text{s}} = -2J\,, \tag{14.2.4}$$

the *exchange energy*. Depending on the overlap, this quantity may be positive as well as negative; in one case the triplet, and in the other, the singlet state of the two spins is energetically more favorable.

When the higher-lying states of the two-atom, two-electron system are ignored because they cannot be excited thermally, and only the lowest-lying singlet and triplet states are considered, the Hamiltonian that determines the state of the electrons in this subspace of the Hilbert space can be replaced by a particularly simple effective Hamiltonian, as it was pointed out by HEISENBERG and DIRAC.[3] By making use of the property that the eigenvalue of the square of the total spin $\boldsymbol{S} = \boldsymbol{s}_1 + \boldsymbol{s}_2$ is zero in the singlet state ($S = 0$) and two in the triplet state ($S = 1$), the effective Hamiltonian

$$\mathcal{H}_{\text{eff}} = \varepsilon_{\text{s}} + \tfrac{1}{2}(\varepsilon_{\text{t}} - \varepsilon_{\text{s}})\boldsymbol{S}^2 = \varepsilon_{\text{s}} - J\boldsymbol{S}^2 \tag{14.2.5}$$

indeed leads to the same energy eigenvalues as the full Hamiltonian in the space of the singlet and triplet states. Since

$$\boldsymbol{S}^2 = (\boldsymbol{s}_1 + \boldsymbol{s}_2)^2 = \boldsymbol{s}_1^2 + \boldsymbol{s}_2^2 + 2\boldsymbol{s}_1 \cdot \boldsymbol{s}_2\,, \tag{14.2.6}$$

and the eigenvalues for $\boldsymbol{s}_1^2$ and $\boldsymbol{s}_2^2$ are both $3/4$, the effective Hamiltonian can be written as

$$\mathcal{H}_{\text{eff}} = -2J\left(\boldsymbol{s}_1 \cdot \boldsymbol{s}_2 + \tfrac{3}{4}\right) + \varepsilon_{\text{s}}\,. \tag{14.2.7}$$

This form of the effective Hamiltonian is called the Heisenberg exchange Hamiltonian (also known as the Heisenberg–Dirac(–Van Vleck) Hamiltonian), and J is the Heisenberg exchange coupling. Considering spins as classical vectors, the triplet state – which is energetically more favorable when $J > 0$ – can be pictured as a state with two parallel spins. This is the situation of ferromagnetic coupling. On the other hand, when $J < 0$ the two spins are antiparallel in the classical picture. This corresponds to an antiferromagnetic coupling.

14.2.2 Indirect Exchange in Metals

Direct exchange is the consequence of the Pauli principle, i.e., the antisymmetry of the total wavefunction of two electrons with spins S_1 and S_2. Direct exchange can occur only if there is a finite overlap between the wavefunctions of the two electrons. This mechanism cannot account for the magnetic properties of every single magnetically ordered material. There are a large number of magnetic materials in which the distance between neighboring magnetic atoms is large, and therefore direct exchange can be ignored. In such materials the spins of core electrons localized on magnetic ions interact via another mechanism in which the other electrons of the system play an essential intermediary role.

[3] W. HEISENBERG, 1926, and P. A. M. DIRAC, 1926.

One possibility that is realized in metals is that the interaction is mediated by mobile conduction electrons. This mechanism can play an important role in elements of the iron group, while it becomes clearly dominant in rare-earth metals. This indirect interaction between two spins will be derived in the Appendix I of Volume 2. Below we shall briefly illustrate the processes that give rise to the interaction, using some concepts, applicable to the more or less freely moving electrons, that will be thoroughly discussed only in Volume 2. Suppose that a localized spin $\boldsymbol{S}_1$ is placed at $\boldsymbol{R}_1 = 0$, and its projection along the quantization axis is S_1^z. If this spin interacts with the conduction electrons through a so-called *s–d* interaction (characterized by the coupling constant $J_{\mathrm{s-d}}$), it will modify the distribution of conduction electrons around itself. The total charge density – which is just the sum of the electron densities for the two spin orientations, $\rho(r) = \rho_\uparrow(r) + \rho_\downarrow(r)$ – remains uniform in space, while the spin density, $\sigma(r) = \frac{1}{2}(\rho_\uparrow(r) - \rho_\downarrow(r))$ becomes position dependent. Far from the localized moment spin density decreases as $1/r^3$ but oscillates with a wavelength determined by the formula $2k_{\mathrm{F}}\lambda = 2\pi$, where k_{F} is the Fermi wave number, the characteristic wave number of the system of electrons. This is the *Ruderman–Kittel oscillation*,[4] which will be discussed in Chapter 29. According to the results derived there,

$$\sigma(r) = \frac{k_{\mathrm{F}}^3}{6\pi^2\hbar^2}\frac{3m_{\mathrm{e}}J_{\mathrm{s-d}}S_1^z k_{\mathrm{F}}}{\pi}F(2k_{\mathrm{F}}r)\,, \tag{14.2.8}$$

where

$$F(x) = \frac{x\cos x - \sin x}{x^4}\,. \tag{14.2.9}$$

Since the spatial periodicity of the oscillation is determined by k_{F}, which depends on the number of free electrons, and is therefore characteristic of the system of electrons, it may be incommensurate with the periodicity of the lattice.

If another localized moment with spin $\boldsymbol{S}_2$ is placed at a distance r from $\boldsymbol{S}_1$, its energetically most favorable orientation is determined by the local value of the spin density induced by $\boldsymbol{S}_1$. When the projection of $\boldsymbol{S}_2$ along the z-axis is S_2^z, its interaction with the local spin density gives a contribution

$$E = -\tfrac{1}{2}J_{\mathrm{s-d}}S_2^z\left[\rho_\uparrow(r) - \rho_\downarrow(r)\right] = S_1^z S_2^z\frac{4m_{\mathrm{e}}J_{\mathrm{s-d}}^2 k_{\mathrm{F}}^4}{(2\pi)^3\hbar^2}F(2k_{\mathrm{F}}r) \tag{14.2.10}$$

to the energy. The same would be obtained if the interaction between the two rigidly fixed spins were assumed to be given by the effective Hamiltonian

$$\mathcal{H}_{\mathrm{eff}} = -2J(\boldsymbol{R}_1 - \boldsymbol{R}_2)S_1^z S_2^z\,, \tag{14.2.11}$$

where

$$J(\boldsymbol{r}) = \frac{2m_{\mathrm{e}}J_{\mathrm{s-d}}^2 k_{\mathrm{F}}^4}{(2\pi)^3\hbar^2}F(2k_{\mathrm{F}}r)\,. \tag{14.2.12}$$

[4] M. A. Ruderman and C. Kittel 1954.

Depending on their distance, the parallel or antiparallel alignment of $\boldsymbol{S}_2$ and $\boldsymbol{S}_1$ would be energetically more favorable.

In reality the orientation of the spins is not fixed; conduction electrons and localized spins may mutually flip each other. Nevertheless the total spin of the two localized moments and the electron system is conserved. Thus in certain processes mediated by electrons the spins of conduction electrons remain eventually unchanged, while the two localized moments are flipped. The two localized spins can then be considered to have flipped each other, just like in the case of direct exchange.

As demonstrated in detail in Appendix I of Volume 2, the effective Hamiltonian eventually takes the form

$$\mathcal{H}_{\text{eff}} = -2J(\boldsymbol{R}_1 - \boldsymbol{R}_2)\boldsymbol{S}_1 \cdot \boldsymbol{S}_2 \,, \tag{14.2.13}$$

with the same $J(\boldsymbol{r})$ as above. This is the *RKKY interaction*.[5] Depending on the separation of the atoms, this interaction may be either ferromagnetic or antiferromagnetic. It gives rise to the formation of long-wavelength spiral structures in rare-earth metals, and also plays an important role in the development of quenched disorder in spin glasses.

14.2.3 Superexchange

In 1934 H. A. Kramers proposed another mechanism that would give rise to an effective exchange interaction between two magnetic atoms when there is no overlap between the wavefunctions of the electrons that are responsible for the magnetic behavior (and so direct exchange is not possible). In this mechanism the interaction is not mediated by the conduction electrons but by the electrons localized on the nonmagnetic atoms between the two interacting magnetic atoms. Now the wavefunctions of nonmagnetic atoms have to overlap with the wavefunctions of the electrons that are responsible for magnetism. This indirect exchange is called *superexchange*. The antiferromagnetism of certain fluorides (MnF_2, FeF_2, CoF_2) and oxides (e.g., MnO) can be interpreted in terms of this interaction.

A more precise description of the interaction was given by P. W. Anderson (1950 and 1959). Below we shall present just a simple illustration. Consider two transition-metal ions (e.g., Mn^{2+}), with a number of oxygen, fluorine, chlorine, or some other similar ions between, and suppose that the wavefunction of their outermost p state overlaps with the wavefunction of the d-electrons of the transition-metal ion.

According to Hund's rules, the ground state of the Mn^{2+} ion – which contains five electrons on its 3d shell ($3d^5$) – is $^6S_{5/2}$, with quantum numbers $S = 5/2$, $L = 0$, $J = 5/2$, while the O^{2-} or F^- ion is in state $^1S_0$ on account of its completely filled $2p^6$ shell, and the total moment of the oxygen vanishes.

[5] M. A. Ruderman and C. Kittel 1954, T. Kasuya 1956, and K. Yosida 1957.

The total spin of a system made up of an oxygen atom in a singlet state and two manganese ions of spin $S = 5/2$ is $0, 1, 2, \ldots$, or 5. Just like in a hydrogen molecule, where the energy is different for the singlet and triplet states, different energies are obtained for the six possible values of the total spin, even though the 3d states do not overlap directly, and the elementary interactions are spin independent. The energy of the system thus depends on the relative orientation of the spins of the two manganese ions.

In contrast to the calculation of the hydrogen molecule, one has to go to the fourth order of perturbation theory. Starting from the initial configuration $Mn^{2+}O^{2-}Mn^{2+}$, one has to allow for the possibility of an electron jumping from the oxygen to a manganese ion, i.e., the appearance of configurations $Mn^{+}O^{-}Mn^{2+}$ and $Mn^{2+}O^{-}Mn^{+}$ as intermediate states.

Instead of a detailed calculation we shall just list the intermediate states that lead to an exchange. Suppose that an electron of the O^{2-} ion jumps to the metal ion on the left-hand side. In this intermediate state the Mn^{+} ion of configuration $3d^6$ is in state ^5D, while the O^{-} ion of configuration $2p^5$ is in state ^2P. Now the O^{-} has a spin of 1/2, and can enter into a two-step direct-exchange process with the Mn^{2+} ion on the right-hand side: the two ions exchange an electron, and the spin of the O^{-} ion may be reversed. In the last step an electron jumps back to the O^{-} ion from the Mn^{+} ion on the left-hand side. Since in the meantime a spin reversal occurred on the oxygen ion, the spin of the rendered electron is opposite to the initial one (with which the whole process started). Thus an effective exchange occurs between the two Mn^{2+} ions. The electron-shell configurations and spin orientations in the initial and final, as well as intermediate states are shown in Fig. 14.8.

Mn^{2+}	$(3d^5)$	↑↑↑↑↑	O^{2-}	$(2p^6)$	↑↑↑↓↓↓	Mn^{2+}	$(3d^5)$	↓↓↓↓↓
Mn^{+}	$(3d^6)$	↑↑↑↑↑↓	O^{-}	$(2p^5)$	↑↑↑↓↓	Mn^{2+}	$(3d^5)$	↓↓↓↓↓
Mn^{+}	$(3d^6)$	↑↑↑↑↑↓	O^{2-}	$(2p^6)$	↑↑↑↓↓↓	Mn^{3+}	$(3d^4)$	↓↓↓↓
Mn^{+}	$(3d^6)$	↑↑↑↑↑↓	O^{-}	$(2p^5)$	↑↑↓↓↓	Mn^{2+}	$(3d^5)$	↓↓↓↓↑
Mn^{2+}	$(3d^5)$	↑↑↑↑↓	O^{2-}	$(2p^6)$	↑↑↑↓↓↓	Mn^{2+}	$(3d^5)$	↓↓↓↓↑

Fig. 14.8. The configuration of electron spins in the initial, three intermediate, and final states of the superexchange interaction

Naturally, there exist other processes in which the spin projection is conserved. By collecting them it can be established that the singlet state of the two spins is the energetically most favorable arrangement. This also means that the two spins are not rigidly fixed in opposite directions but can flip each other. Neglecting those states (of higher energy) where the magnitude of the spin is changed, the spectrum of the low-lying states can be identified with that of an effective Heisenberg exchange Hamiltonian. However, the

strength of the superexchange interaction is rather difficult to calculate from first principles. The sign of the coupling depends on the population of the levels. In line with the previous results for the manganese ion, superexchange is antiferromagnetic for the heaviest 3d metals. The antiferromagnetic structure of insulating metal oxides can be interpreted along these lines, and the same kind of superexchange is responsible for the magnetic coupling in high-T_c superconductors that contain Cu–O planes.

14.2.4 Double Exchange

In 1951 C. Zener proposed another exchange mechanism in oxides of mixed valence where an oxygen ion is surrounded by two transition-metal ions of unequal valence. Let us consider the configuration in Fig. 14.9, where the O^{2-} ion is between a Mn^{4+} ion (which has three electrons on its 3d shell) and an Mn^{3+} ion (which has four electrons on its 3d shell).

Mn^{4+}	O^{2-}	Mn^{3+}		Mn^{3+}	O^{2-}	Mn^{4+}
$3d^3$	$2p^6$	$3d^4$	$\Rightarrow$	$3d^4$	$2p^6$	$3d^3$
$S = 3/2$	$S = 0$	$S = 2$		$S = 2$	$S = 0$	$S = 3/2$

Fig. 14.9. The electron configuration of the ions in the initial and final states of the double exchange interaction

Via two subsequent processes, in which an electron jumps from the oxygen to the Mn^{4+} ion, and then the empty level on the oxygen is filled by an electron of the Mn^{3+} ion, an electron of the Mn^{3+} is eventually transferred to the Mn^{4+} ion. In reality, this process can occur only when the spins of the two ions are parallel. To understand this let us examine how the atomic levels are populated by 3d electrons in Mn^{3+} and Mn^{4+} ions. In a cubic crystal field d levels are known to be split into threefold degenerate t_{2g} and twofold degenerate e_g levels. According to Hund's first rule, the spins of the four electrons of the Mn^{3+} ion and of the three electrons of the Mn^{4+} ion are parallel on these levels. Consequently the total electron spin of the Mn^{3+} ion is $S = 2$, while that of the Mn^{4+} ion is $S = 3/2$. This is shown in Fig. 14.10.

When only the initial and final states of the two-step transfer process are considered, it is as if the two manganese ions had exchanged their spins. In this phenomenon, called *double exchange*, an electron from the e_g level of the Mn^{3+} ion hops via the oxygen ion on the so far empty e_g level of the Mn^{4+} ion. According to Hund's first rule the spin of the e_g electron has to be aligned with the spins of the three electrons on the t_{2g} level, otherwise the energy of the state would be much higher. Thus it is energetically unfavorable for a spin-up e_g electron to hop on a Mn^{4+} ion if its total spin $S = 3/2$ points downward. Therefore spin exchange (or the hopping of the electron)

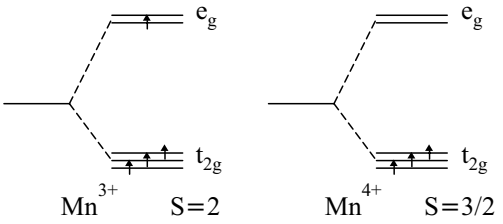


Fig. 14.10. The d electrons of Mn^{3+} and Mn^{4+} ions on the t_{2g} and e_g levels

takes place with high probability only when the spins of neighboring ions are aligned ferromagnetically. Since hopping lowers the total energy by reducing the kinetic energy, ferromagnetic alignment of the spins is energetically favored. Thus, in contrast to superexchange, which is often antiferromagnetic, double exchange is always ferromagnetic. The propagation (hopping) of the electrons in the crystal indicates that such systems are metallic.

Such a double exchange plays an important role in magnetite, which contains divalent (Fe^{2+}, $3d^6$) and trivalent (Fe^{3+}, $3d^5$) iron ions. The spins of the Fe^{2+} and Fe^{3+} ions at octahedral sites are aligned by double exchange interactions. On the other hand, the spins of the Fe^{3+} ions at tetrahedral sites are antiparallel to the spins at octahedral sites because of antiferromagnetic superexchange interactions.

14.3 Simple Models of Magnetism

In the previous section we saw that, up to an additive constant, the interaction between two localized spins $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$ can always be written as

$$\mathcal{H} = -2J\boldsymbol{S}_1 \cdot \boldsymbol{S}_2 \,, \tag{14.3.1}$$

independently of the specific mechanism of the exchange – although the coupling strength J depends on the character of the exchange and on the separation of atomic spins. We now have to generalize this formula to the case when atomic spins are arranged in a regular crystalline structure.

14.3.1 The Isotropic Heisenberg Model

It is not easy to generalize the exchange interaction between two spins to a solid that contains a large number of magnetic atoms. For simplicity we shall assume that a net spin $\boldsymbol{S}_i$ is located at each lattice point $\boldsymbol{R}_i$ of the crystal. It is known that the atomic magnetic moment can also come from orbital motion; indeed, one should therefore deal with the total angular momentum $\boldsymbol{J}_i$ and the corresponding magnetic moment $\boldsymbol{\mu}_i = -g_J\mu_B\boldsymbol{J}_i$. However, the microscopic origin of the moments is of little importance for the phenomena

discussed here, therefore we shall continue to speak about spins and use the notation $\boldsymbol{S}_i$ – although by this we shall invariably mean the total localized angular momentum of the ion. The only notational hint for this will be using g instead of g_{e} or g_J for the g-factor.

The net spin $\boldsymbol{S}_i$ can be the resultant of several electronic spins, and electrons may be exchanged with the electrons of several neighboring atoms. The problem is greatly simplified by the assumption that the total interaction can be written as the sum of pairwise interactions. If it is also assumed that for each atom it is sufficient to consider only the ground-state configuration as determined by Hund's rules (or by the splitting due to the crystal field), and higher-lying excited states can be neglected, then using some simple considerations the results obtained for two electrons are fairly straightforwardly generalized without a full quantum mechanical treatment. In the subspace of the Hilbert space that contains the lowest-lying states the magnitude of each atomic spin (and therefore of each atomic magnetic moment) is constant. Interactions may at most change their orientations. If the interaction between the atomic spins at lattice points $\boldsymbol{R}_i$ and $\boldsymbol{R}_j$ is given by a term of the form of the Heisenberg exchange interaction, $\boldsymbol{S}_i \cdot \boldsymbol{S}_j$, and the strength of the interaction can be characterized by the exchange coupling – i.e., the amplitude $J_{ij} = J(\boldsymbol{R}_i - \boldsymbol{R}_j)$, which depends on the separation of the spins –, then the system of spins is described by the effective Hamiltonian

$$\boxed{\mathcal{H} = -\sum_{ij} J_{ij} \boldsymbol{S}_i \cdot \boldsymbol{S}_j \,.} \tag{14.3.2}$$

Since the indices i and j both run over the entire lattice, each pair of spins occurs with the correct weight factor. It was first pointed out by W. Heisenberg (1928), and independently of him by J. Frenkel that ferromagnetism can be explained in terms of such pairwise interactions between localized spins. For this reason the model described by the above Hamiltonian is called the *Heisenberg model* of magnetism. According to the foregoing, states in which electrons are transferred from one atom to the other (for no matter how short a period) are not present explicitly in this model. Only the orientation of atomic moments is considered.

At low temperatures the exchange interaction does not usually leave atomic spins independent: they tend to align each other into a preferred direction, giving rise to an ordered structure. The simplest of them is *ferromagnetic order*.This is observed in materials in which the sign of the exchange interaction is positive, hence it lines up neighboring spins in the same direction. However, the exchange interaction is not necessarily of ferromagnetic character. As it was discussed at the beginning of this chapter, ferromagnets form a relatively small group of magnetically ordered materials. More numerous are antiferromagnets, in which the dominant nearest-neighbor coupling is antiferromagnetic ($J_{ij} < 0$), which gives rise to an ordered array of alternately

oriented spins – the details of which depend on the coupling between more distant neighbors.

In magnetic materials described with the isotropic Heisenberg model the interaction fixes only the relative orientation of spins; it does not determine their exact spatial orientation, as the Hamiltonian itself possesses full spherical symmetry (O(3) or SU(2) symmetry). The ordered ferromagnetic or antiferromagnetic state does not show this continuous spherical symmetry. The continuous symmetry of the Hamiltonian is therefore spontaneously broken in the actual state. This has important consequences on the low-energy excitation spectrum of the system. Namely, according to Goldstone's theorem presented in Chapter 6, when the ground state of the system breaks a continuous symmetry of the Hamiltonian, there must be a gapless bosonic branch in the excitation spectrum.

Turning on a symmetry-breaking magnetic field facilitates the mathematical determination of the symmetry-breaking solutions of the Hamiltonian. Since magnetization will point in this direction, it is practical to choose this direction as the spin quantization axis. At the end of the calculation we may take the limit of vanishing magnetic field. In this approach the Hamiltonian is customarily written as

$$\mathcal{H} = -\sum_{i,j} J_{ij} \boldsymbol{S}_i \cdot \boldsymbol{S}_j - g\mu_{\mathrm{B}} \boldsymbol{B} \cdot \Big(\sum_i \boldsymbol{S}_i \Big) , \tag{14.3.3}$$

where g is the g-factor of the localized moment. If the moment comes entirely from the spin then g is the g-factor of the electron, which is negative. If the total atomic moment has a part that comes from the orbital angular momentum then the Landé factor g_J must be used, however comparison with (3.2.70) shows that $g = -g_J$.

The field $\boldsymbol{B}$ can be replaced by $\mu_0 \boldsymbol{H}$. If we were to use the full magnetic induction $\mu_{\mathrm{r}} \mu_0 \boldsymbol{H}$ for $\boldsymbol{B}$ then the effects of other moments would be counted twice in magnetically ordered materials as the exchange part of the Heisenberg model contains precisely the effects of the neighbors.

14.3.2 Anisotropic Models

We have seen that the effective Hamiltonian of the exchange interaction between two spins is spherically symmetric, therefore it is invariant under arbitrary rotations of the two spins through a common angle. For spins in a crystal lattice the interaction is expected to depend on the orientation of the spins with respect to the crystalline axes, since the symmetries of the crystal also have to appear in the form of spin–spin interactions. When only terms bilinear in the spin are allowed, the magnetic behavior for systems with cubic symmetry remains to be described by the Hamiltonian

$$\mathcal{H} = -\sum_{i,j} J_{ij} \boldsymbol{S}_i \cdot \boldsymbol{S}_j , \tag{14.3.4}$$

since the only bilinear expression in the spin components that is invariant under the symmetries of a cubic crystal is the isotropic combination

$$\mathcal{H} = -\sum_{i,j} J_{ij} \left\{ S_i^x S_j^x + S_i^y S_j^y + S_i^z S_j^z \right\} . \tag{14.3.5}$$

In tetragonal crystals one axis has a privileged status over the two others. This uniaxial anisotropy has to be reflected in the magnetic Hamiltonian, too. Choosing the z-axis along this direction, the Hamiltonian that possesses the symmetries of the crystal is

$$\mathcal{H} = -\sum_{i,j} J_{ij} \left\{ S_i^x S_j^x + S_i^y S_j^y + \Delta S_i^z S_j^z \right\} . \tag{14.3.6}$$

When $\Delta > 1$, the z components of the spins are coupled most strongly, so a nonvanishing magnetization appears preferentially along this direction. In this case the magnetic material has an easy axis of magnetization.When $\Delta < 1$, the spin components in the (x, y) plane are coupled strongly, and so magnetization appears in this plane – the easy plane of magnetization. In an extreme case of uniaxial anisotropy only the z components of the spins are coupled:

$$\mathcal{H}_{\text{Ising}} = -\sum_{i,j} J_{ij} S_i^z S_j^z . \tag{14.3.7}$$

This is the Ising model,[6] a fundamental model of statistical physics, since the partition function can be calculated as a sum over all possible classical configurations of the spin projections. On the other hand, when $\Delta = 0$, the XY model is recovered:

$$\mathcal{H}_{\text{XY}} = -\sum_{i,j} J_{ij} \left\{ S_i^x S_j^x + S_i^y S_j^y \right\} . \tag{14.3.8}$$

In tetragonal crystals, in addition to the contribution of ion pairs to uniaxial anisotropy, a term that corresponds to single-ion anisotropy,

$$-D \sum_i (S_i^z)^2 \tag{14.3.9}$$

may also appear in the Hamiltonian. In cubic crystals the first anisotropic contribution contains the fourth-order product of spins. The single-ion anisotropy term is

$$\mathcal{H}_{\text{aniso}} = -K \sum_i \left[(S_i^x)^4 + (S_i^y)^4 + (S_i^z)^4 \right] . \tag{14.3.10}$$

We shall choose the z-axis as the quantization direction, and instead of the operators S^x and S^y we shall repeatedly use the spin-projection-lowering and -raising operators

[6] E. Ising, 1925. It would be more appropriate to call it the Lenz–Ising model, since it was proposed by Ising's supervisor, W. Lenz, in 1920.

$$S_i^+ = S_i^x + \mathrm{i}S_i^y\,, \qquad S_i^- = S_i^x - \mathrm{i}S_i^y\,. \tag{14.3.11}$$

In terms of these the Hamiltonian of the isotropic and uniaxially anisotropic systems are

$$\mathcal{H} = -\sum_{i,j} J_{ij}\left\{\tfrac{1}{2}\left(S_i^+ S_j^- + S_i^- S_j^+\right) + S_i^z S_j^z\right\} \tag{14.3.12}$$

and

$$\mathcal{H} = -\sum_{i,j} J_{ij}\left\{\tfrac{1}{2}\left(S_i^+ S_j^- + S_i^- S_j^+\right) + \Delta S_i^z S_j^z\right\}. \tag{14.3.13}$$

These forms show even more clearly that the Heisenberg model – unlike the Ising model – is truly quantum mechanical. And although the classical model is more easily treated, we shall deal with the quantum mechanical model, as it corresponds more closely to physical reality.

14.4 The Mean-Field Approximation

In the introductory section magnetically ordered states were characterized by the relative orientation of a classical vector, the expectation value of the spin or magnetic moment. At low temperatures the orientation of spins is such that their projection along the preferred direction is maximum with high probability. As temperature is increased, the expectation value of the magnetic moment decreases and finally vanishes. The simplest description of this phenomenon can be given in the framework of the mean-field theory. This approximation is based on the assumption that the alignment of individual spins is not sensitive to the thermal or quantum fluctuations of neighboring spins. Assuming that neighboring moments can be replaced by their average, they create an effective static magnetic field, and the magnetic moment of the atom lines up with this. On the other hand, when the g-factor in $\boldsymbol{\mu}_i = g\mu_\mathrm{B}\boldsymbol{S}_i$ (i.e., the negative of the Landé factor) is negative, spins line up in the opposite direction.

To determine the field generated by the neighbors, we shall write the spin operator $\boldsymbol{S}_i$ in the equivalent form

$$\boldsymbol{S}_i = \langle \boldsymbol{S}_i\rangle + \left(\boldsymbol{S}_i - \langle \boldsymbol{S}_i\rangle\right)\,, \tag{14.4.1}$$

where $\langle\rangle$ denotes the thermal average. Substituting this in the Hamiltonian,

$$\begin{aligned}\mathcal{H} &= -\sum_{i,j} J_{ij}\left(\boldsymbol{S}_i - \langle \boldsymbol{S}_i\rangle + \langle \boldsymbol{S}_i\rangle\right)\cdot\left(\boldsymbol{S}_j - \langle \boldsymbol{S}_j\rangle + \langle \boldsymbol{S}_j\rangle\right) - g\mu_0\mu_\mathrm{B}\boldsymbol{H}\cdot\Big(\sum_i \boldsymbol{S}_i\Big)\\ &= \sum_{i,j} J_{ij}\langle \boldsymbol{S}_i\rangle\langle \boldsymbol{S}_j\rangle - 2\sum_{i,j} J_{ij}\langle \boldsymbol{S}_j\rangle\boldsymbol{S}_i - g\mu_0\mu_\mathrm{B}\boldsymbol{H}\cdot\Big(\sum_i \boldsymbol{S}_i\Big)\\ &\quad -\sum_{i,j} J_{ij}\left(\boldsymbol{S}_i - \langle \boldsymbol{S}_i\rangle\right)\cdot\left(\boldsymbol{S}_j - \langle \boldsymbol{S}_j\rangle\right).\end{aligned} \tag{14.4.2}$$

The mean-field approximation consists of neglecting the last term that is of second order in the fluctuations, i.e., in the deviations $\boldsymbol{S}_i - \langle \boldsymbol{S}_i \rangle$ from the mean value. The Hamiltonian obtained in this way is linear in the spin operators, so it can be rewritten as

$$\mathcal{H}_{\text{mean field}} = \sum_{i,j} J_{ij} \langle \boldsymbol{S}_i \rangle \langle \boldsymbol{S}_j \rangle - g\mu_{\text{B}} \boldsymbol{B}_{\text{eff}} \cdot \Big(\sum_i \boldsymbol{S}_i \Big) , \tag{14.4.3}$$

where

$$\boldsymbol{B}_{\text{eff}} = \mu_0 \boldsymbol{H} + \frac{2}{g\mu_{\text{B}}} \sum_j J_{ij} \langle \boldsymbol{S}_j \rangle . \tag{14.4.4}$$

Using the magnetic moment instead of the spin,

$$\boldsymbol{B}_{\text{eff}} = \mu_0 \boldsymbol{H} + \frac{2}{g^2 \mu_{\text{B}}^2} \sum_j J_{ij} \langle \boldsymbol{\mu}_j \rangle . \tag{14.4.5}$$

Even in the absence of an external magnetic field, the effective magnetic field – which depends on the average magnetic moment of the neighboring spins – can give rise to magnetic ordering. This effective internal field is called the *mean field* or the *Weiss field*, as it was P. Weiss (1907) who gave the first phenomenological account of ferromagnetism based on the assumption of such an internal field. The term *molecular field* is also used.

14.4.1 The Mean-Field Theory of Ferromagnetism

In a ferromagnetically ordered state the mean value of the atomic magnetic moment is the same at each lattice point, therefore $\boldsymbol{B}_{\text{eff}}$ is independent of the choice of the lattice point. Making use of the relation $\boldsymbol{M} = N\langle \boldsymbol{\mu}_i \rangle / V$ between magnetization and atomic moment,

$$\boldsymbol{B}_{\text{eff}} = \mu_0 \boldsymbol{H} + \frac{2V}{Ng^2 \mu_{\text{B}}^2} \sum_j J_{ij} \boldsymbol{M} . \tag{14.4.6}$$

As in the ferromagnetic case ($J_{ij} > 0$) the internal field is parallel to the external field, scalar quantities can be used:

$$B_{\text{eff}} = \mu_0 H + \frac{2V}{Ng^2 \mu_{\text{B}}^2} \sum_j J_{ij} M . \tag{14.4.7}$$

In the previous formula the exchange interaction appears in the combination

$$J_0 = \sum_j J_{ij} . \tag{14.4.8}$$

Expressing the effective field in terms of this,

$$B_{\text{eff}} = \mu_0 H + \lambda M\,, \qquad \text{where} \qquad \lambda = \frac{2V}{N g^2 \mu_{\text{B}}^2} J_0\,. \tag{14.4.9}$$

The coefficient λ specifies the relation between the internal field and magnetization in the Weiss theory. If exchange occurs only with the z nearest neighbors then $J_0 = Jz$. When second- and third-neighbor interactions cannot be neglected, J_0 is the weighted average of the various coupling constants. The necessary condition for ferromagnetic ordering is that J_0 should be positive. This can happen even if some of the latter couplings are antiferromagnetic.

It was derived on page 57 in Section 3.2.6 on paramagnetism that the magnetization due to the magnetic moment of an atom of spin S in a magnetic field B is

$$M = \frac{N}{V} |g| \mu_{\text{B}} S B_S(\beta |g| \mu_{\text{B}} S B)\,, \tag{14.4.10}$$

where $B_S(x)$ is the Brillouin function. In the mean-field theory this expression has to be modified by replacing the magnetic induction B by the effective field B_{eff}, leading to

$$M = \frac{N}{V} |g| \mu_{\text{B}} S B_S(\beta |g| \mu_{\text{B}} S B_{\text{eff}})\,. \tag{14.4.11}$$

In contrast to the paramagnetic case, the effective field itself also depends on magnetization. Therefore the self-consistent equation

$$M = \frac{N}{V} |g| \mu_{\text{B}} S B_S\big(\beta |g| \mu_{\text{B}} S (\mu_0 H + \lambda M)\big) \tag{14.4.12}$$

has to be solved now.

In the absence of an external magnetic field the argument of the Brillouin function is

$$x = \beta |g| \mu_{\text{B}} S \lambda M = \frac{2VSJ_0 M}{N |g| \mu_{\text{B}} k_{\text{B}} T}\,. \tag{14.4.13}$$

Expressing condition (14.4.12) of self-consistency in terms of this variable,

$$\frac{k_{\text{B}} T}{2 J_0 S^2} x = B_S(x)\,. \tag{14.4.14}$$

The solutions to this equation are easily obtained graphically by plotting the left- and right-hand sides separately, as in Fig. 14.11. The thermal average of the magnetic moment can be read off from the intersection points.

The slope of the expression on the left-hand side is proportional to temperature. Therefore above a certain critical temperature T_{C} only the trivial solution $x = 0$ – and consequently $M = 0$ – exists. The same solution exists also below T_{C}, however it becomes unstable. In this regime there exists another solution with finite magnetization whose free energy is lower.

According to the expansion (3.2.84), the slope of the Brillouin function at $x = 0$ is $(S+1)/3S$, thus the equation that determines the critical temperature of the ferromagnetic state, the *Curie temperature* is

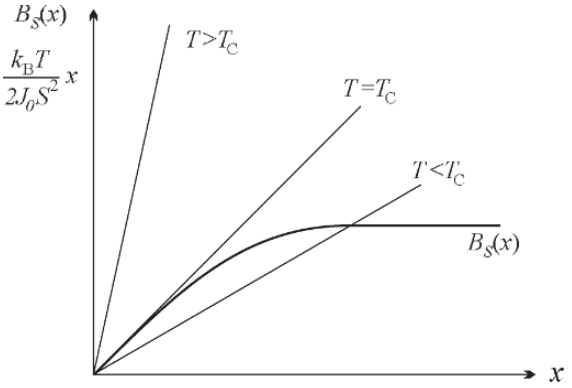


Fig. 14.11. Graphical determination of the mean-field magnetization from the intersection points of the Brillouin function and straight lines of temperature-dependent slope

$$\frac{k_B T_C}{2J_0 S^2} = \frac{S+1}{3S}\,, \tag{14.4.15}$$

from which

$$\boxed{T_C = \frac{2J_0 S(S+1)}{3k_B}\,.} \tag{14.4.16}$$

Spontaneous magnetization vanishes above T_C. However, an external magnetic field aligns the magnetic moments, at least partially. The spin system then behaves paramagnetically. To determine the magnetic susceptibility, we shall take (14.4.11) and (14.4.12), but this time in the presence of a magnetic field. For weak fields it is sufficient to keep the leading, linear term in the expansion of the right-hand side, obtained via (3.2.84). In this approximation

$$\begin{aligned} M &= \frac{N}{V}|g|\mu_B S \frac{S+1}{3S}\frac{|g|\mu_B}{k_B T} S B_{\text{eff}} \\ &= \frac{N}{V}|g|\mu_B S \frac{S+1}{3S}\frac{|g|\mu_B}{k_B T} S \left(\mu_0 H + \frac{2V}{N g^2 \mu_B^2} J_0\, M\right). \end{aligned} \tag{14.4.17}$$

Substituting T_C from the boxed formula,

$$M = \frac{N}{V}\frac{g^2\mu_B^2}{k_B T}\frac{S(S+1)}{3}\mu_0 H + \frac{k_B T_C}{k_B T} M\,. \tag{14.4.18}$$

Solving this equation for magnetization,

$$M = \frac{N}{V}\frac{g^2\mu_B^2 S(S+1)}{3k_B(T-T_C)}\mu_0 H\,. \tag{14.4.19}$$

Thus in the paramagnetic region the susceptibility satisfies the *Curie–Weiss law*:

$$\boxed{\chi = \frac{N}{V}\frac{g^2\mu_{\mathrm{B}}^2\mu_0 S(S+1)}{3k_{\mathrm{B}}(T-T_{\mathrm{C}})}\,.} \tag{14.4.20}$$

It is immediately seen that the susceptibility diverges as the critical temperature T_{C} of the ferromagnetic transition is approached.

Below but close to T_{C} an explicit formula can be obtained for the temperature dependence of spontaneous magnetization as well by taking into account the first correction beyond the linear term in the expansion of the Brillouin function. The self-consistent solution of the equation then leads to

$$M \propto (T_{\mathrm{C}} - T)^{1/2}\,. \tag{14.4.21}$$

In the critical point itself magnetization is a nonlinear function of the magnetic field, as (14.4.18) shows that the term which is linear in magnetization drops out at $T = T_{\mathrm{C}}$. This leads to

$$M \propto H^{1/3}\,. \tag{14.4.22}$$

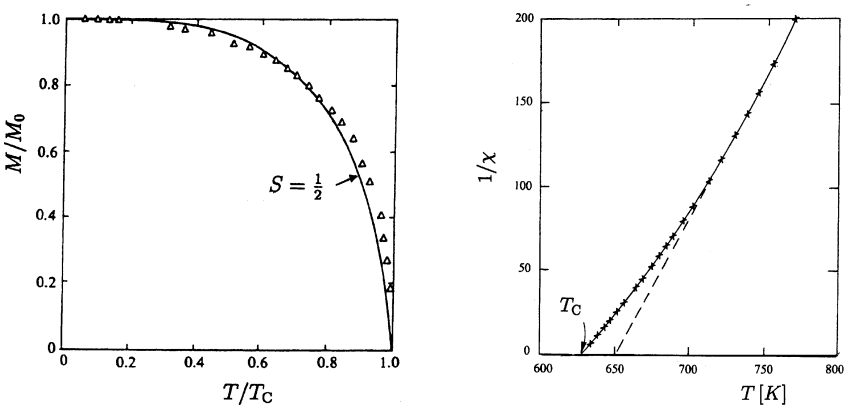


Fig. 14.12. Comparison of the measured magnetic properties of nickel [P. Weiss and R. Forrer, *Ann. de Phys.* **5**, 153 (1926)] with the results obtained in the mean-field theory for $S = 1/2$. (*a*) Magnetization and (*b*) inverse susceptibility, as functions of temperature

At low temperatures magnetization tends to the saturation value $M_0 = (N/V)|g|\mu_{\mathrm{B}}S$. Using the expansion $\coth x \approx 1 + 2\mathrm{e}^{-2x}$ in the form (3.2.79) of the Brillouin function for large values of its argument, the leading-order contribution is

$$B_S(x) \approx 1 - \frac{1}{S}\mathrm{e}^{-x/S}\,. \tag{14.4.23}$$

In this region the first correction to the temperature dependence of magnetization is obtained by replacing the magnetization by its saturation value on the right-hand side of (14.4.12). Then

$$M = M_0 \left[1 - \frac{1}{S} \exp\left(-\frac{3}{S+1} \frac{T_\mathrm{C}}{T} \right) \right] . \tag{14.4.24}$$

At intermediate temperatures no analytical expression can be derived. Figure 14.12 shows the measured temperature dependence of magnetization and inverse susceptibility for nickel, compared with the results obtained in the mean-field approximation for a spin $S = 1/2$ using the Brillouin function, and from the Curie–Weiss law.

A similar behavior is observed in other ferromagnetic materials, too. This leads to the conclusion that the mean-field theory provides a qualitatively good description for the finite-temperature behavior of the properties of ferromagnetic materials, however, its quantitative predictions are incorrect at low temperatures as well as close to the critical point – where analytical results can be obtained. To improve the description, one has to go beyond the mean-field approximation. At low temperatures a more accurate treatment of the low-energy excited states is required, while in the vicinity of the critical point critical fluctuations need to be taken into account.

14.4.2 The Mean-Field Theory of Antiferromagnetism

We have seen that even in cubic crystals various antiferromagnetic structures may occur. The reason for this is that besides the exchange interaction between nearest-neighbor moments, second- and third-neighbor couplings also play a role in determining the relative orientation of magnetic moments. This can be demonstrated most easily on the example of structures formed in a simple cubic lattice.

We shall consider the spins $\boldsymbol{S}_i$ localized at lattice points as classical vectors, expand the spin density

$$\boldsymbol{S}(\boldsymbol{r}) = \sum_i \boldsymbol{S}_i \delta(\boldsymbol{r} - \boldsymbol{R}_i) \tag{14.4.25}$$

into a Fourier series, and assume that the magnetic structure can be represented by a single Fourier component, i.e.,

$$\boldsymbol{S}(\boldsymbol{r}) = \boldsymbol{S}_{\boldsymbol{q}} \mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r}} , \tag{14.4.26}$$

which is equivalent to saying that localized spins can be given in the form

$$\boldsymbol{S}_i = \boldsymbol{S}_{\boldsymbol{q}} \mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{R}_i} . \tag{14.4.27}$$

This choice is justified for collinear structures, where the upward or downward direction of the spin vector is determined by the phase factors. Taking the

magnitude of the vector $\boldsymbol{S_q}$ to be S, and assuming that the strength of the exchange interaction is J_1 between first and J_2 between second neighbors, the energy per spin is

$$\begin{aligned} E/N = &-2J_1S^2\big[\cos q_xa + \cos q_ya + \cos q_za\big] \\ &-2J_2S^2\big[\cos(q_x+q_y)a + \cos(q_x-q_y)a + \cos(q_y+q_z)a \\ &\quad + \cos(q_y-q_z)a + \cos(q_z+q_x)a + \cos(q_z-q_x)a\big]\,. \end{aligned} \tag{14.4.28}$$

The energy expression takes its minima for vectors $\boldsymbol{q}$ whose components are related by

$$\sin q_xa = \sin q_ya = \sin q_za = 0 \tag{14.4.29}$$

or by

$$\cos q_xa = \cos q_ya = \cos q_za = -J_1/4J_2\,, \tag{14.4.30}$$

provided $|J_2| > |J_1|/4$. The latter would lead to an incommensurate structure, however the energy associated with such a structure is never an absolute minimum. Apart from the vector $\boldsymbol{q}_0 = (0,0,0)$, which corresponds to the ferromagnetic state, the inequivalent solutions of interest correspond to the vectors

$$\boldsymbol{q}_1 = (\pi/a)(0,0,1), \quad \boldsymbol{q}_2 = (\pi/a)(1,1,0), \quad \boldsymbol{q}_3 = (\pi/a)(1,1,1)\,. \tag{14.4.31}$$

These are just the vectors given in (14.1.3); they correspond to the structures shown in Fig. 14.2. From the minimum of the energy it can be determined which structure is stable for given signs and ratio of J_1 and J_2. This is shown in Fig. 14.13.

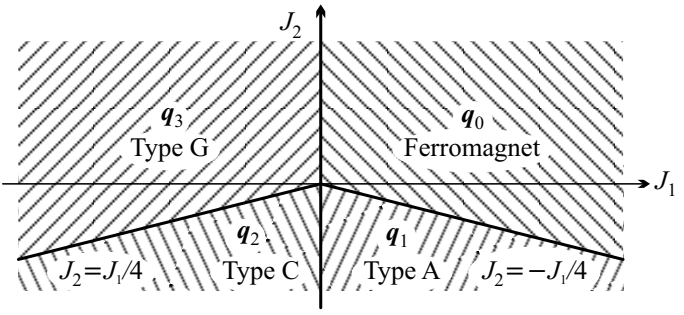


Fig. 14.13. The range of stability for magnetic structures in a simple cubic crystal associated with different vectors $\boldsymbol{q}$ in the space of the couplings J_1 and J_2

The stability condition for various antiferromagnetic structures can be determined along the same lines in other crystal lattices, too.

In some of the presented antiferromagnetic structures upward- and downward-pointing atomic spins are arranged in such a way as to make up

two simple interpenetrating sublattices in such a way that the nearest neighbors of each atom with an up spin are atoms with a down spin, and vice versa. This situation is observed in the antiferromagnetic structure formed in a simple cubic lattice – shown in Fig. 14.2(c) – in which up and down spins alternate along the edges, forming face-centered cubic sublattices. Another example that occurs in body-centered cubic lattices is shown in Fig. 14.3(a); here the spins at the vertices of the cubic primitive cell point in the opposite direction as spins at the cell centers. Up and down spins now make up simple cubic sublattices. Other antiferromagnetic structures, especially those formed in face-centered cubic lattices, are more complicated. Whichever atom is chosen, among its neighbors there are some with parallel and some with antiparallel spin, and so the decomposition requires at least four sublattices: two with spins up, and two with spins down.

For simplicity, we shall deal only with the properties of antiferromagnetic materials that can be decomposed into two sublattices, and in which each of the N atoms in volume V is surrounded by z first neighbors located on the other sublattice. We shall denote the sublattices by A and B, and their lattice points by i and j, respectively. We shall also assume that the exchange interaction acts only between nearest neighbors: $J_{ij} = J < 0$, where i and j denote nearest neighbor lattice points, which are necessarily on opposite sublattices. The Hamiltonian can be written as

$$\mathcal{H} = -2 \sum_{\langle i,j \rangle} J \boldsymbol{S}_i \cdot \boldsymbol{S}_j - g\mu_{\mathrm{B}} \mu_0 \boldsymbol{H} \cdot \Big(\sum_{i \in A} \boldsymbol{S}_i + \sum_{j \in B} \boldsymbol{S}_j \Big) , \tag{14.4.32}$$

where label i runs over sublattice A, and label j over sublattice B, and $\langle i, j \rangle$ in the first term denotes the constraint that i and j should be nearest neighbors. Since each pair occurs only once in the sum, a factor of two appears in front of the first term compared to the ferromagnetic case.

The mean-field approximation is introduced in the same way as for ferromagnets. By neglecting second-order terms in the deviation from the mean value,

$$\begin{aligned} \mathcal{H}_{\text{mean field}} = {} & 2 \sum_{\langle i,j \rangle} J \langle \boldsymbol{S}_i \rangle \langle \boldsymbol{S}_j \rangle - 2 \sum_{\langle i,j \rangle} J \langle \boldsymbol{S}_i \rangle \boldsymbol{S}_j - 2 \sum_{\langle i,j \rangle} J \langle \boldsymbol{S}_j \rangle \boldsymbol{S}_i \\ & - g\mu_{\mathrm{B}} \mu_0 \boldsymbol{H} \cdot \Big(\sum_i \boldsymbol{S}_i + \sum_j \boldsymbol{S}_j \Big) . \end{aligned} \tag{14.4.33}$$

This expression can be interpreted again by saying that spins feel an effective field $\boldsymbol{B}_{\text{eff}}$. However spins on opposite sublattices are aligned in different directions, so the effective fields are different on the two sublattices – but they are independent of the lattice point on each sublattice:

$$\boldsymbol{B}_{\text{eff, A}} = \mu_0 \boldsymbol{H} + \frac{2}{g\mu_\text{B}} \sum_{j \in B} J \langle \boldsymbol{S}_j \rangle \,, \tag{14.4.34-a}$$

$$\boldsymbol{B}_{\text{eff, B}} = \mu_0 \boldsymbol{H} + \frac{2}{g\mu_\text{B}} \sum_{i \in A} J \langle \boldsymbol{S}_i \rangle \,. \tag{14.4.34-b}$$

Along with spins, magnetic moments will also be different on the two sublattices. Let us therefore introduce the magnetization of the sublattices by the definition

$$\boldsymbol{M}_\text{A} = \frac{1}{V} g\mu_\text{B} \sum_{i \in A} \langle \boldsymbol{S}_i \rangle \,, \qquad \boldsymbol{M}_\text{B} = \frac{1}{V} g\mu_\text{B} \sum_{j \in B} \langle \boldsymbol{S}_j \rangle \,. \tag{14.4.35}$$

Sublattice magnetizations are aligned with the effective field on the sublattice, and their magnitude is given by

$$\begin{aligned} M_\text{A} &= \frac{N}{2V} |g| \mu_\text{B} S B_S(\beta |g| \mu_\text{B} S B_{\text{eff, A}}) \,, \\ M_\text{B} &= \frac{N}{2V} |g| \mu_\text{B} S B_S(\beta |g| \mu_\text{B} S B_{\text{eff, B}}) \,, \end{aligned} \tag{14.4.36}$$

in analogy with (14.4.12). The factor 1/2 appeared because each sublattice contains $N/2$ magnetic atoms.

We shall first examine the ordered phase, in which sublattice magnetizations are along an axis, one of them pointing upward, and the other downward: $M_\text{A} = -M_\text{B}$. Owing to the antiferromagnetic coupling $J < 0$, in the absence of an external magnetic field the effective field on sublattice A can also be written as

$$B_{\text{eff, A}} = \frac{4V}{Ng^2\mu_\text{B}^2} z J M_\text{B} = \frac{4V}{Ng^2\mu_\text{B}^2} z |J| M_\text{A} \,. \tag{14.4.37}$$

Substituting this into the self-consistent equation for the sublattice magnetization, the same form is obtained as in the ferromagnetic case – with the only difference that the number of atoms is now $N/2$ instead of N. Sublattice magnetization depends on temperature in the same way as spontaneous magnetization of ferromagnets does. Following the same steps as in the mean-field theory of ferromagnetism, the critical temperature of magnetic ordering – called the *Néel temperature* – is given by

$$\boxed{T_\text{N} = \frac{2|J|zS(S+1)}{3k_\text{B}} \,.} \tag{14.4.38}$$

To determine the behavior of susceptibility above T_N, we take the linear expansion of the Brillouin function in the presence of a magnetic field:

$$\begin{aligned} M_\text{A} &= \frac{N}{2V} \frac{g^2\mu_\text{B}^2}{k_\text{B}T} S^2 \frac{S+1}{3S} \left(\mu_0 H + \frac{4V}{Ng^2\mu_\text{B}^2} J z M_\text{B} \right) , \\ M_\text{B} &= \frac{N}{2V} \frac{g^2\mu_\text{B}^2}{k_\text{B}T} S^2 \frac{S+1}{3S} \left(\mu_0 H + \frac{4V}{Ng^2\mu_\text{B}^2} J z M_\text{A} \right) . \end{aligned} \tag{14.4.39}$$

The sum of the two equations gives

$$M_{\mathrm{A}} + M_{\mathrm{B}} = \frac{N}{V}\frac{g^2\mu_{\mathrm{B}}^2\mu_0 S(S+1)}{3k_{\mathrm{B}}T}H + \frac{2JzS(S+1)}{3k_{\mathrm{B}}T}\left(M_{\mathrm{A}} + M_{\mathrm{B}}\right). \quad (14.4.40)$$

Making use of formula (14.4.38) for the Néel temperature,

$$M_{\mathrm{A}} + M_{\mathrm{B}} = \frac{N}{V}\frac{g^2\mu_{\mathrm{B}}^2\mu_0 S(S+1)}{3k_{\mathrm{B}}T}H - \frac{k_{\mathrm{B}}T_{\mathrm{N}}}{3k_{\mathrm{B}}T}\left(M_{\mathrm{A}} + M_{\mathrm{B}}\right). \quad (14.4.41)$$

Solving this equation for the net magnetization,

$$M = M_{\mathrm{A}} + M_{\mathrm{B}} = \frac{N}{V}\frac{g^2\mu_{\mathrm{B}}^2\mu_0 S(S+1)}{3k_{\mathrm{B}}(T+T_{\mathrm{N}})}H\,. \quad (14.4.42)$$

This leads to the following temperature dependence of the magnetic susceptibility in the paramagnetic phase above the Néel temperature:

$$\boxed{\chi = \frac{N}{V}\frac{g^2\mu_{\mathrm{B}}^2\mu_0 S(S+1)}{3k_{\mathrm{B}}(T+T_{\mathrm{N}})}\,.} \quad (14.4.43)$$

This expression is analogous to the Curie–Weiss law, which gives the susceptibility of ferromagnets above the Curie temperature; the only difference is that T_{C} is now replaced by $-T_{\mathrm{N}}$. For this reason, susceptibility remains finite at the critical temperature of the magnetic transition, that is to say the divergence observed in the Curie point of ferromagnets does not appear. This might be surprising at first, since both the ferromagnetic and antiferromagnetic phase transitions are of second order, and fluctuations are expected to play equally important roles at the critical temperature in the two cases. However, the order parameters are different, so the critical behavior manifests itself in different quantities.

In ferromagnets the magnetic order parameter is the uniform magnetization. Its long-wavelength fluctuations govern the critical behavior around the critical point T_{C}, and give rise to the divergence of susceptibility in a uniform magnetic field. In contrast, the order parameter in antiferromagnets is the sublattice magnetization, more precisely $M_{\mathrm{A}} - M_{\mathrm{B}}$. According to the Fourier representation (14.1.2), when the magnetic moment varies periodically in space this difference is proportional to the Fourier coefficient associated with the wave vector $\boldsymbol{q}_0$ of the antiferromagnetic structure. Therefore fluctuations do not become critical in the long-wavelength limit either but at wavelengths that correspond to the periodicity of the antiferromagnetic structure. By considering the wave-number-dependent susceptibility – i.e., the magnetic response to finite-wavelength perturbations – as a function of $\boldsymbol{k}$, no divergencies are observed at $\boldsymbol{k} = 0$ (which is associated with a uniform field), only around $\boldsymbol{k} = \boldsymbol{q}_0$. This component of the susceptibility corresponds to the response to an alternating magnetic field (staggered field) that is opposite on the two sublattices. It is therefore called *staggered susceptibility*.

Below the Néel temperature the net magnetization vanishes. If the orientations of the moments on the two sublattices would remain each other's opposite in an external magnetic field, energy could be reduced only by making the lengths of the moments unequal on the two sublattices. In fact in the energetically most favorable arrangement the sublattice magnetizations are not aligned with the field lines: they are tilted symmetrically and make an angle θ with the direction perpendicular to the field, as shown in Fig. 14.14(a).

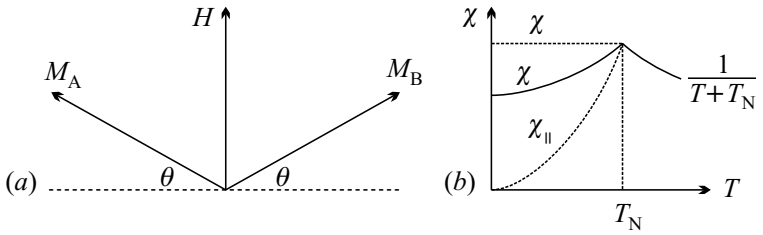


Fig. 14.14. (a) The orientation of the sublattice-magnetization vectors of an isotropic antiferromagnet in a magnetic field. (b) The temperature dependence of susceptibility

In the mean-field theory the sublattice magnetization $\boldsymbol{M}_{\mathrm{A}}$ of sublattice A must be aligned with the effective field acting on the sublattice,

$$\boldsymbol{B}_{\mathrm{eff,\,A}} = \mu_0 \boldsymbol{H} + \frac{4V}{Ng^2\mu_{\mathrm{B}}^2} Jz \boldsymbol{M}_{\mathrm{B}} = \mu_0 \boldsymbol{H} + \lambda \boldsymbol{M}_{\mathrm{B}}\,, \tag{14.4.44}$$

that is, just like $\boldsymbol{M}_{\mathrm{A}}$, $\boldsymbol{B}_{\mathrm{eff,\,A}}$ also has to make an angle θ with the direction perpendicular to the magnetic field $\boldsymbol{H}$. Writing this condition in terms of the components of $\boldsymbol{B}_{\mathrm{eff,\,A}}$, and bearing in mind that the Weiss coefficient λ is now negative,

$$\frac{\mu_0 H - |\lambda| M_{\mathrm{B}} \sin\theta}{|\lambda| M_{\mathrm{B}} \cos\theta} = \tan\theta\,. \tag{14.4.45}$$

In weak magnetic fields θ is small, so the moments are almost perpendicular to the field. Expanding the trigonometric functions to leading order,

$$\mu_0 H = 2|\lambda| M_{\mathrm{B}} \theta\,. \tag{14.4.46}$$

A similar expression applies to the magnetization of the other sublattice, too. Expressing the component of the induced magnetization in the direction of the magnetic field,

$$M = (M_{\mathrm{A}} + M_{\mathrm{B}}) \sin\theta \approx (M_{\mathrm{A}} + M_{\mathrm{B}})\,\theta = \frac{\mu_0 H}{|\lambda|}\,. \tag{14.4.47}$$

Substitution of the expression for λ into this formula gives the susceptibility, which is independent of the sublattice magnetization and therefore of temperature as well:

$$\chi_\perp = \frac{N}{V}\frac{g^2\mu_B^2\mu_0}{4|J|z} = \frac{N}{V}\frac{g^2\mu_B^2\mu_0 S(S+1)}{6k_B T_N}\,. \tag{14.4.48}$$

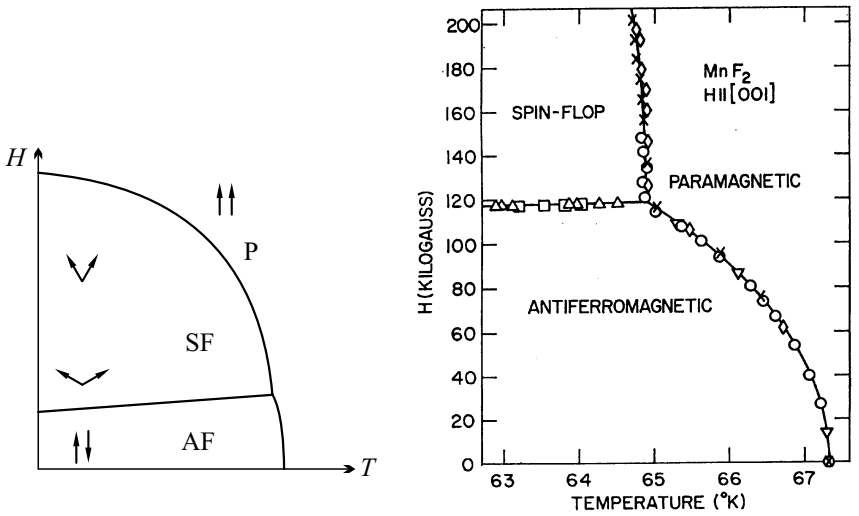
The symbol $\perp$ in the subscript refers to the almost perpendicular orientation of the moments with respect to the magnetic field.

Comparing this result with (14.4.43), the expression for the susceptibility above the Néel temperature, the same value is obtained at the critical point. Thus in an isotropic antiferromagnet susceptibility is constant below the Néel temperature, and starts to decrease above it. As illustrated in Fig. 14.14(b), at very high temperatures the same $1/T$ dependence is observed as in paramagnets (see Curie's law on page 52).

However, there are no perfectly isotropic materials in which this behavior could be observed. The inevitable anisotropy stemming from the crystalline structure gives rise to interactions that will keep spins along a preferred crystallographic direction. If the weak external magnetic field also acts in this direction, then anisotropy does not let spins rotate away from this direction and be almost perpendicular to the field. The susceptibility measured under such circumstances is the so-called *parallel susceptibility*, $\chi_\parallel$. To determine it, one should start with (14.4.36), but assume that the magnetization of one sublattice is along the external magnetic field, while that of the other is along the opposite direction. At finite temperatures, where the magnitude of the sublattice magnetization is smaller than the saturation value, the magnetic field increases the magnetization on the sublattice where it is aligned with the field, and decreases it on the other sublattice. The result is a highly temperature dependent net magnetization and susceptibility – which vanish in the $T \to 0$ limit. The temperature dependence of parallel susceptibility is shown in Fig. 14.14(b). In macroscopic samples the direction of sublattice magnetization may vary from domain to domain. Therefore measurements of the susceptibility give a weighted average of $\chi_\parallel$ and $\chi_\perp$, specifically

$$\chi = \frac{1}{3}\chi_\parallel + \frac{2}{3}\chi_\perp\,. \tag{14.4.49}$$

According to the foregoing, in antiferromagnets that exhibit sufficiently strong uniaxial anisotropy sublattice magnetization is along the easy axis of magnetization even in the presence of a magnetic field. Spins will rotate away from this direction – that is, the magnetic field will overcome anisotropy – only when the magnetic field applied in the direction of easy magnetization exceeds a threshold value. In such a field spins will suddenly turn from the parallel direction to a configuration in which the sublattice magnetizations are oriented symmetrically and make an angle θ with the easy axis of magnetization. This first-order phase transition is therefore called *spin-flop transition*, and the resulting phase is the *spin-flop phase*. When the field is increased even further, the sublattice magnetizations become more and more aligned with the field, and finally a perfectly collinear structure arises through a second-order phase transition. The theoretically predicted and experimentally measured phase diagrams are shown in Fig. 14.15.



(*a*) (*b*)

Fig. 14.15. The phase diagram (temperature vs. magnetic field plot) of an anisotropic antiferromagnet with an easy axis of magnetization. Part (*a*) shows a schematic phase diagram, while part (*b*) is the measured phase diagram of MnF_2 below the Néel temperature [Y. Shapira and S. Foner, *Phys. Rev. B* **1**, 3083 (1970)]

14.4.3 The General Description of Two-Sublattice Antiferromagnets

The above simple treatment of antiferromagnets led to the result that in the paramagnetic region the magnetic susceptibility can be written as

$$\chi = \frac{C}{T + T_{\mathrm{N}}}\,, \tag{14.4.50}$$

where T_{N} is the Néel temperature. Experiments, on the other hand, have shown that while in the vast majority of antiferromagnetic materials the temperature dependence of susceptibility is fairly well approximated by the formula

$$\chi = \frac{C}{T + \Theta}\,, \tag{14.4.51}$$

the value of Θ is nevertheless different from the Néel temperature. As data in Table 14.6 indicate, the deviation can sometimes be very large.

This deviation is the result of the applied approximation, whereby only the exchange interaction between nearest neighbors was taken into account, and so the behavior of each sublattice was entirely determined by the effective field due to the other sublattice. To demonstrate this, we shall determine the critical temperature and the Θ parameter of two-sublattice antiferromagnets

Table 14.6. Néel temperature and the characteristic temperature Θ that appears in the fit of susceptibility data

Material	T_{N}(K)	Θ(K)	Material	T_{N}(K)	Θ(K)
MnO	118	610	FeS	593	917
FeO	185	570	$KMnF_3$	88	158
CoO	291	330	MnF_2	72	113
NiO	518	2000	Cr_2O_3	318	1070

for the case when the exchange interaction between second neighbors (located on the same sublattice) cannot be neglected.

We shall denote the strength of the exchange interaction between nearest neighbors on opposite sublattices by J_{AB}; the number of nearest neighbors (on sublattice B) of an atom on sublattice A by z_{AB}, and the same for an atom on sublattice B by z_{BA}. Along the same lines, we shall denote the strength of the exchange interaction between second neighbors (located on the same sublattice) by J_{AA} and J_{BB}, and the number of second neighbors by z_{AA} and z_{BB}, respectively. We shall also assume that the two sublattices are of the same structure, that is $z_{\mathrm{AB}} = z_{\mathrm{BA}}$, $J_{\mathrm{AA}} = J_{\mathrm{BB}}$, and $z_{\mathrm{AA}} = z_{\mathrm{BB}}$. In the mean-field theory the sublattice magnetization has to be determined self-consistently from the equations

$$\begin{aligned} M_{\mathrm{A}} &= \frac{N}{2V}|g|\mu_{\mathrm{B}}SB_S(\beta|g|\mu_{\mathrm{B}}SB_{\mathrm{eff,\,A}})\,, \\ M_{\mathrm{B}} &= \frac{N}{2V}|g|\mu_{\mathrm{B}}SB_S(\beta|g|\mu_{\mathrm{B}}SB_{\mathrm{eff,\,B}})\,, \end{aligned} \tag{14.4.52}$$

where the effective field acting on the spins of sublattice A is given by the generalization of (14.4.37),

$$B_{\mathrm{eff,\,A}} = \frac{4V}{Ng^2\mu_{\mathrm{B}}^2}z_{\mathrm{AB}}J_{\mathrm{AB}}M_{\mathrm{B}} + \frac{4V}{Ng^2\mu_{\mathrm{B}}^2}z_{\mathrm{AA}}J_{\mathrm{AA}}M_{\mathrm{A}}\,. \tag{14.4.53}$$

The effective field acting on sublattice B is given by an analogous formula.

In the absence of a magnetic field, and in the vicinity of the phase transition point, where the sublattice magnetization is small, the expansion of the Brillouin function leads to the system of equations

$$\begin{aligned} M_{\mathrm{A}} &= \frac{2S(S+1)}{3k_{\mathrm{B}}T}\left(z_{\mathrm{AB}}J_{\mathrm{AB}}M_{\mathrm{B}} + z_{\mathrm{AA}}J_{\mathrm{AA}}M_{\mathrm{A}}\right), \\ M_{\mathrm{B}} &= \frac{2S(S+1)}{3k_{\mathrm{B}}T}\left(z_{\mathrm{BA}}J_{\mathrm{BA}}M_{\mathrm{A}} + z_{\mathrm{BB}}J_{\mathrm{BB}}M_{\mathrm{B}}\right). \end{aligned} \tag{14.4.54}$$

Nontrivial solutions exist only when the determinant of the coefficients vanishes. When the nearest-neighbor interaction is antiferromagnetic, i.e., $J_{\mathrm{AB}} < 0$, the physically meaningful solution for the transition temperature is

$$T_{\mathrm{N}} = \frac{2S(S+1)}{3k_{\mathrm{B}}} \left(z_{\mathrm{AB}} |J_{\mathrm{AB}}| + z_{\mathrm{AA}} J_{\mathrm{AA}} \right) . \tag{14.4.55}$$

To determine the susceptibility in the disordered phase, a weak external field is turned on. The effective field on the atoms in sublattice A is then

$$B_{\mathrm{eff,\,A}} = \mu_0 H + \frac{4V}{N g^2 \mu_{\mathrm{B}}^2} z_{\mathrm{AB}} J_{\mathrm{AB}} M_{\mathrm{B}} + \frac{4V}{N g^2 \mu_{\mathrm{B}}^2} z_{\mathrm{AA}} J_{\mathrm{AA}} M_{\mathrm{A}} . \tag{14.4.56}$$

Using the linear expansion of the Brillouin function, the sublattice magnetization in this field is

$$M_{\mathrm{A}} = \frac{N}{2V} g^2 \mu_{\mathrm{B}}^2 \frac{S(S+1)}{3k_{\mathrm{B}}T} \left[\mu_0 H + \frac{4V}{N g^2 \mu_{\mathrm{B}}^2} (J_{\mathrm{AB}} z_{\mathrm{AB}} M_{\mathrm{B}} + J_{\mathrm{AA}} z_{\mathrm{AA}} M_{\mathrm{A}}) \right] . \tag{14.4.57}$$

Adding this expression to the analogous one for the other sublattice,

$$\begin{aligned} M_{\mathrm{A}} + M_{\mathrm{B}} = {} & \frac{N}{V} \frac{g^2 \mu_{\mathrm{B}}^2 \mu_0 S(S+1)}{3k_{\mathrm{B}}T} H \\ & + \frac{2S(S+1)}{3k_{\mathrm{B}}T} \left(J_{\mathrm{AB}} z_{\mathrm{AB}} + J_{\mathrm{AA}} z_{\mathrm{AA}} \right) \left(M_{\mathrm{A}} + M_{\mathrm{B}} \right) . \end{aligned} \tag{14.4.58}$$

Using the form (14.4.51) for the susceptibility, some rearrangement leads to

$$\Theta = \frac{2S(S+1)}{3k_{\mathrm{B}}} \left(z_{\mathrm{AB}} |J_{\mathrm{AB}}| - z_{\mathrm{AA}} J_{\mathrm{AA}} \right) . \tag{14.4.59}$$

Comparison with expression (14.4.55) for the Néel temperature shows that $\Theta > T_{\mathrm{N}}$ if the second-neighbor interaction is also antiferromagnetic. This applies to the majority of cases.

14.4.4 The Mean-Field Theory of Ferrimagnetism

Within the framework of the mean-field theory, ferrimagnets can be treated much in the same manner as antiferromagnets. In the simplest case the lattice may be decomposed into two sublattices, however, the sublattice magnetizations do not cancel each other now. Either because the magnetic atoms on the two sublattices are unequal in number, or because their magnetic moments are of unequal magnitude. We shall examine the latter case, and generalize the equations derived for the sublattice magnetizations in antiferromagnets by letting the spin magnitudes (S_{A} and S_{B}) as well as the exchange integrals be different on the two sublattices:

$$\begin{aligned} M_{\mathrm{A}} &= \frac{N}{2V} |g| \mu_{\mathrm{B}} S_{\mathrm{A}} B_{S_{\mathrm{A}}} (\beta |g| \mu_{\mathrm{B}} S_{\mathrm{A}} B_{\mathrm{eff,\,A}}) , \\ M_{\mathrm{B}} &= \frac{N}{2V} |g| \mu_{\mathrm{B}} S_{\mathrm{B}} B_{S_{\mathrm{B}}} (\beta |g| \mu_{\mathrm{B}} S_{\mathrm{B}} B_{\mathrm{eff,\,B}}) . \end{aligned} \tag{14.4.60}$$

In the presence of an applied magnetic field, the effective fields on the two sublattices are

$$\begin{aligned} B_{\mathrm{eff,\,A}} &= \mu_0 H + \lambda_{\mathrm{AA}} M_{\mathrm{A}} + \lambda_{\mathrm{AB}} M_{\mathrm{B}} \,, \\ B_{\mathrm{eff,\,B}} &= \mu_0 H + \lambda_{\mathrm{BA}} M_{\mathrm{A}} + \lambda_{\mathrm{BB}} M_{\mathrm{B}} \,. \end{aligned} \tag{14.4.61}$$

Depending on the signs and magnitudes of the coupling constants, the temperature dependence of the net magnetization traces out strikingly different curves, as shown in Fig. 14.16.

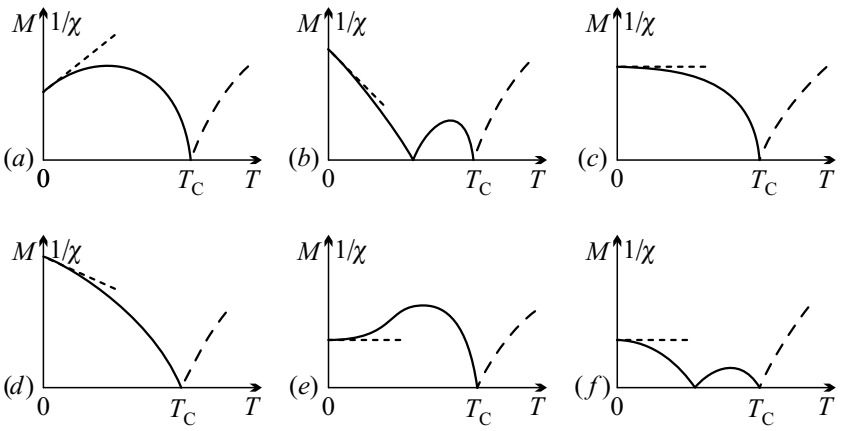


Fig. 14.16. Various possibilities for the temperature dependence of magnetization in ferrimagnets. Dashed lines indicate the slope at $T = 0$. Above T_{C} long-dashed lines show the inverse susceptibility

In the simplest case the temperature dependence of the full magnetization is similar to that of ferromagnets. At low temperatures variations are very small, until the magnetization gradually decreases to zero at a finite temperature T_{C}. In some materials magnetization may nevertheless vary substantially with temperature even at very low temperatures, while in others magnetization may not be a monotonic function of temperature but have a maximum instead. Even more interesting is the case where the oppositely directed sublattice magnetizations cancel (compensate) each other at a finite temperature $T_{\mathrm{comp}} < T_{\mathrm{C}}$. The net magnetization vanishes at this temperature, and reappears at higher temperatures. This is shown for gadolinium–iron garnet in Fig. 14.7.

14.5 The General Description of Magnetic Phase Transitions

The low-temperature magnetically ordered phase is always separated from the high-temperature paramagnetic phase by a first- or second-order phase

transition. In the previous section this transition was described using the Heisenberg Hamiltonian, and the temperature dependence of magnetization and susceptibility were determined on both sides of the transition point. In all our previous examples the transitions were of second order, that is magnetization and sublattice magnetization appeared gradually, not with an abrupt discontinuity. While this is not always the case, second-order magnetic phase transitions are very common, therefore we shall give a concise overview of the general theory of second-order transitions. We shall demonstrate that the characteristic physical quantities of the magnetic system can be derived without making specific assumptions about the Hamiltonian, and that the effects of fluctuations may also be taken into account by going beyond the mean-field approximation.

14.5.1 The Landau Theory of Second-Order Phase Transitions

In 1937 L. D. Landau[7] put forward a general phenomenological theory for the description of second-order phase transitions. He considered the sudden change of symmetry as one of the most important characteristics of such transitions. At high temperatures the system is in a disordered state that possesses high symmetry. At lower temperatures a more ordered – and consequently less symmetric – state occurs, therefore the higher symmetry of the previous phase is broken. The two phases are distinguished by the existence and nonexistence of some symmetries, therefore a sharp line can be drawn between the two. In Landau's approach order can be characterized quantitatively by an order parameter, which is zero in the disordered phase and nonzero in the ordered phase. The physical meaning of the order parameter is determined by the character of the transition. In ferromagnetic ordering magnetization itself can serve as an order parameter, while in antiferromagnetic ordering the sublattice magnetization is a good candidate. In order–disorder transitions in alloys the concentration of individual components on the sublattices is a possible choice. When superfluidity and superconductivity are analyzed, the wavefunction of the condensate proves to be an appropriate order parameter. Below we shall demonstrate that, regardless of the physical meaning of the order parameter, general statements can be made about the variations of the free energy and other physical quantities during the phase transition.

Landau based the description of second-order phase transitions on the following assumptions:

1. The phase transition is continuous. This means that the order parameter ψ_0 – which vanishes in the disordered phase but takes a nonzero value in the ordered phase below the critical temperature T_c – varies continuously in the phase transition. If the order parameter changed discontinuously to a finite value, the phase transition would be of first order.

[7] See footnote on page 28.

2. A free-energy-like quantity F can be introduced, which can be defined even for the nonequilibrium values ψ of the order parameter. This quantity has its minimum at the equilibrium value ψ_0 of the order parameter, and this minimum value is just the equilibrium value of the free energy. Therefore F must have its minimum at $\psi = 0$ in the disordered phase and at a finite ψ_0 in the ordered phase (i.e., below the critical temperature T_c).
3. F is an analytic function of ψ, that is, it can be expanded into a power series of ψ. Since the order parameter varies continuously in the vicinity of the transition point, only the first few terms need to be retained if the behavior of the system is studied in a small temperature range around T_c.
4. The expansion coefficients are analytical functions of temperature.

For simplicity we shall assume that the order parameter is a real scalar, although this is not always the case.[8] It follows from the above assumptions that the expansion of F is

$$F = F_0 + A(T)\psi^2 + \tfrac{1}{2}B(T)\psi^4 + \dots . \tag{14.5.1}$$

The linear and cubic terms are missing for simple reasons. The linear term must be absent, otherwise ψ_0, the order parameter at the minimum of the free energy, would be nonzero everywhere – with the possible exception of the point where the coefficient of the linear term vanishes –, and so there would be no order–disorder transition. The cubic term must also be absent, otherwise the position of the minimum would not change continuously with temperature – recall that the temperature dependence appears through the coefficients A and B – but ψ_0 would jump abruptly from zero to a finite value, which would correspond to a first-order phase transition. Generally speaking, the requirement that the cubic term should vanish imposes severe restrictions on the possible transitions. In magnetic transitions the situation is simplified right from the start by the symmetry properties of the magnetic moment that permit only even powers of ψ. The free energy is plotted against ψ in Fig. 14.17; the minimum is at $\psi = 0$ on one graph, and at a finite value ψ_0 on the other.

The previous expression has its minimum at $\psi = 0$ if the coefficients A and B are both positive. The minimum is at a finite ψ_0 if $A < 0$ and $B > 0$. Therefore the following temperature dependence is required for the coefficients:

$$A(T) \begin{cases} > 0\,, & \text{if} \quad T > T_c\,, \\ < 0\,, & \text{if} \quad T < T_c\,, \end{cases} \tag{14.5.2}$$

while $B(T)$ has to be positive in both regions. Assuming that the coefficients are analytic, the simplest choice is

$$A(T) = a(T - T_c) \qquad \text{and} \qquad B(T) = B(T_c) = B > 0\,. \tag{14.5.3}$$

[8] Generalization to an n-component order parameter is straightforward.

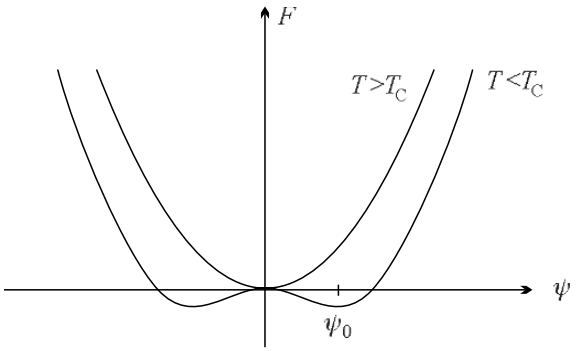


Fig. 14.17. Free energy as a function of the order parameter in the vicinity of the critical temperature

Determined from the minimum condition for the free energy, the equilibrium value of the order parameter is

$$\psi_0^2(T) = \begin{cases} 0\,, & \text{if} \quad T \geq T_\mathrm{c}\,, \\ -\dfrac{A(T)}{B}\,, & \text{if} \quad T < T_\mathrm{c}\,. \end{cases} \tag{14.5.4}$$

The order parameter appears continuously in T_c indeed; its temperature dependence is

$$\psi_0(T) \sim \sqrt{T_\mathrm{c} - T}\,. \tag{14.5.5}$$

Using the terminology of critical phenomena: the critical exponent of the order parameter is $\beta = 1/2$ in the Landau theory.

This value is used in the determination of the equilibrium value of the free energy F in the ordered phase. F is smaller than the free energy F_0 in the disordered phase:

$$F = F_0 - \frac{A^2(T)}{2B} = F_0 - \frac{a^2}{2B}(T - T_\mathrm{c})^2\,. \tag{14.5.6}$$

It follows immediately that the specific heat is not continuous but has a finite jump in the phase transition point:

$$\Delta C = \frac{a^2}{B} T_\mathrm{c}\,. \tag{14.5.7}$$

If the order parameter is linearly coupled to some external field, e.g., magnetization to the external magnetic field, then the susceptibility associated with the response to the external field can also be determined from the Landau theory. In magnetic systems the energy contribution $-\boldsymbol{M} \cdot \boldsymbol{B}$ due to the magnetic field has to be taken into account. Choosing the order parameter as

the component of magnetization along the direction of the field, the equilibrium value of the magnetization can be determined from the minimum of the free energy

$$F = F_0 + A(T)M^2 + \tfrac{1}{2}BM^4 - \mu_0 MH\,, \tag{14.5.8}$$

which leads to

$$2A(T)M + 2BM^3 = \mu_0 H\,. \tag{14.5.9}$$

Keeping only the leading-order term above the critical temperature T_c,

$$M = \frac{\mu_0 H}{2A(T)} = \frac{\mu_0 H}{2a(T - T_\mathrm{c})}\,, \tag{14.5.10}$$

so the magnetic susceptibility is

$$\chi_\mathrm{m} = \frac{\mu_0}{2a(T - T_\mathrm{c})}\,. \tag{14.5.11}$$

The same temperature dependence has appeared as in the Curie–Weiss law for ferromagnets. In antiferromagnets, where sublattice magnetization is the order parameter, susceptibility measured in a uniform field remains finite, however the staggered susceptibility shows the same kind of singularity at $\boldsymbol{k} = \boldsymbol{q}_0$, the characteristic wave vector of the antiferromagnetic structure.

The differential susceptibility in the ordered phase can be determined from the relation

$$2A(T)\frac{\partial M}{\partial H} + 6BM^2\frac{\partial M}{\partial H} = \mu_0\,, \tag{14.5.12}$$

which is derived from (14.5.9). Substituting the order parameter from (14.5.4),

$$\chi_\mathrm{m} = -\frac{\mu_0}{4A(T)} = \frac{\mu_0}{4a(T_\mathrm{c} - T)}\,. \tag{14.5.13}$$

Susceptibility diverges as the -1st power of $|T - T_\mathrm{c}|$ on both sides of the transition point, thus the critical exponent is $\gamma = 1$.

14.5.2 Determination of Possible Magnetic Structures

Up to now the order parameter has been assumed to be a real scalar, therefore only one second- and one fourth-order term appeared in the expansion of the free energy. As we shall see in Chapter 26, the analysis of superconductivity requires that a complex order parameter be used. However, even in magnetic systems the order parameter may be a multicomponent quantity. One could think that the vector character of magnetization requires a three-component order parameter. But the situation is not so simple. The number of components and the form of the second- and fourth-order terms in the free energy expansion are determined by the symmetries of the system.

The free energy of a crystal has to be invariant under all the symmetry operations that take the crystal into itself. To obtain such an invariant expression

for the free energy when the symmetries of the disordered phase are known, the order parameter has to be expanded in terms of the basis functions of the irreducible representations of the symmetry group of the disordered phase, and then second- and fourth-order invariants have to be constructed from the expansion coefficients. For simplicity, we shall assume that there are two irreducible representations: a two- and a three-dimensional one with normalized basis functions $\phi_i^{(1)}$ and $\phi_i^{(2)}$. The expansion of the order parameter is then

$$\psi = (c_1\phi_1^{(1)} + c_2\phi_2^{(1)}) + (d_1\phi_1^{(2)} + d_2\phi_2^{(2)} + d_3\phi_3^{(2)})\,. \tag{14.5.14}$$

Transferring the transformation properties of basis functions to the coefficients, only those combinations may appear in the free energy that are invariant under all symmetries of the high-temperature phase. A single second-order invariant is associated with each irreducible representation, and these appear with independent coefficients in the expansion of the free energy:

$$F = F_0 + A_1(T)(c_1^2 + c_2^2) + A_2(T)(d_1^2 + d_2^2 + d_3^2) + \ldots\,. \tag{14.5.15}$$

We shall start from the disordered phase, where each coefficient $A_i(T)$ is positive, otherwise the minimum of F would not occur at vanishing order parameter. In the ordered phase those coefficients (c_i, d_j, or others if there are further irreducible representations) acquire a finite value that are associated with that particular irreducible representation for which the temperature-dependent coefficients A_i disappears first (i.e., at the highest temperature). Therefore the dimension of the order parameter is determined by the dimension of the corresponding irreducible representation.

Recall that the irreducible representations of the crystal's space group are specified by the wave vector $\boldsymbol{k}$, which characterizes the behavior of the system under spatial translations, and more than one irreducible representation may be associated with a vector $\boldsymbol{k}$. The coefficient $A_i(T)$ obviously depends on the wave vector $\boldsymbol{k}$ associated with the irreducible representation. To emphasize this dependence, the notation $A_{i\boldsymbol{k}}(T)$ will be used. According to the previous considerations, the transition temperature T_c and the wave vector $\boldsymbol{k}_0$ characterizing the behavior of the magnetic structure under translations are determined from the condition that it is the highest temperature where one of the coefficients $A_{i\boldsymbol{k}}(T)$ vanishes. We shall demonstrate that barring accidental solutions (which are not the consequences of symmetry considerations) $\boldsymbol{k}_0$ must be a high-symmetry point of the Brillouin zone.

This is because when the coefficients are continuous functions of $\boldsymbol{k}$ there is no reason why the linear term of the expansion around a general $\boldsymbol{k}_0$ in powers of $\boldsymbol{k} - \boldsymbol{k}_0$ should be missing. Consequently, if the coefficient at wave vector $\boldsymbol{k}_0$ vanishes at a temperature T_c, i.e., $A_{i\boldsymbol{k}_0}(T_\mathrm{c}) = 0$, then the linear expansion of the coefficients in both $\boldsymbol{k} - \boldsymbol{k}_0$ and $T - T_\mathrm{c}$ shows that it vanishes at a higher temperature for some wave vector $\boldsymbol{k}$ that is different from $\boldsymbol{k}_0$ – in striking contradiction with our assumption. This argument does not apply to high-symmetry points of the Brillouin zone: symmetry considerations may

imply that the coefficient has an extremum at such a $\boldsymbol{k}_0$. This explains why in the overwhelming majority of magnetic structures the wave vector $\boldsymbol{k}_0$, which describes the spatial modulation of the ordered magnetic moment, is a high-symmetry point of the Brillouin zone.

14.5.3 Spatial Inhomogeneities and the Correlation Length

The Landau theory leads to the same results as the mean-field theory since fluctuations are neglected in both. However, experiments show deviations from the Curie–Weiss law, and magnetization does not vanish in the Curie point with the exponent predicted by the mean-field theory, as illustrated in Fig. 14.12. This stems from the assumption that the order parameter is perfectly homogeneous in space. In reality, spatial inhomogeneities may become important in the vicinity of the phase transition point. Owing to thermal fluctuations, order is not perfectly homogeneous under T_c, and even above T_c the order parameter may acquire a nonzero value locally, over small regions. Nevertheless we cannot speak of a magnetically ordered state, since order can be observed only over a finite correlation length ξ – in other words, only short-range order exists, and there is no correlation between more distant regions. To account for spatial inhomogeneities, we shall write the full free energy of the system in terms of a free-energy density $f(\boldsymbol{r})$:

$$F = \int f(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r}\,, \tag{14.5.16}$$

where $f(\boldsymbol{r})$ can be expressed in a form similar to (14.5.1) using a position-dependent local order parameter $\psi(\boldsymbol{r})$. Because of the spatial variations, a new contribution may appear that is proportional to the square of the gradient of the order parameter and that corresponds to the energy increment due to spatial inhomogeneities:

$$f(\boldsymbol{r}) = f_0 + A(T)\psi^2(\boldsymbol{r}) + \tfrac{1}{2}B(T)\psi^4(\boldsymbol{r}) + \tfrac{1}{2}C(T)\left(\boldsymbol{\nabla}\psi(\boldsymbol{r})\right)^2 + \ldots\,. \tag{14.5.17}$$

If the order parameter ψ fluctuates about ψ_0, the order parameter of the homogeneous system satisfying the equation

$$2A(T)\psi_0 + 2B(T)\psi_0^3 = 0\,, \tag{14.5.18}$$

then it can be written as $\psi(\boldsymbol{r}) = \psi_0 + \delta\psi(\boldsymbol{r})$, where $\delta\psi(\boldsymbol{r})$ is small and is subject to the constraint

$$\int \delta\psi(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r} = 0\,. \tag{14.5.19}$$

Using this decomposition, the leading-order increment of the free energy due to fluctuations is

$$\delta F = \int \mathrm{d}\boldsymbol{r}\Big[\tfrac{1}{2}C\big(\boldsymbol{\nabla}\delta\psi(\boldsymbol{r})\big)^2 + A(T)\big(\delta\psi(\boldsymbol{r})\big)^2 + 3B(T)\psi_0^2\big(\delta\psi(\boldsymbol{r})\big)^2\Big]. \tag{14.5.20}$$

Expanding fluctuations into a Fourier series,

$$\delta\psi(\boldsymbol{r}) = \frac{1}{V}\sum_{\boldsymbol{k}} \psi_{\boldsymbol{k}}\, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\,, \tag{14.5.21}$$

where the homogeneous term that corresponds to $\boldsymbol{k} = 0$ is missing, thus the contribution of the fluctuations to the free energy is

$$\delta F = \frac{1}{V}\sum_{\boldsymbol{k}} \left[\tfrac{1}{2}C\boldsymbol{k}^2 + A(T) + 3B(T)\psi_0^2\right]|\psi_{\boldsymbol{k}}|^2\,. \tag{14.5.22}$$

Since the probability for such a fluctuation to occur is given by the Boltzmann factor

$$P(\delta F) = \exp\left(-\frac{\delta F}{k_{\mathrm{B}}T}\right)\,, \tag{14.5.23}$$

the thermal average of the square of the component $\psi_{\boldsymbol{k}}$ is

$$\left\langle|\psi_{\boldsymbol{k}}|^2\right\rangle = \frac{k_{\mathrm{B}}TV}{C\boldsymbol{k}^2 + 2A(T) + 6B(T)\psi_0^2}\,. \tag{14.5.24}$$

Above T_{c}, where $\psi_0 = 0$ the amplitude of long-wavelength fluctuations increases as the temperature approaches the transition point. The same is true below T_{c}, where $\psi_0^2 = -A(T)/B$, that is

$$\left\langle|\psi_{\boldsymbol{k}}|^2\right\rangle = \frac{k_{\mathrm{B}}TV}{C\boldsymbol{k}^2 - 4A(T)}\,. \tag{14.5.25}$$

Writing this expression in the form

$$\left\langle|\psi_{\boldsymbol{k}}|^2\right\rangle = \frac{\chi}{1 + (k\xi)^2}\,, \tag{14.5.26}$$

the *correlation length* ξ is defined as

$$\xi^2 = \begin{cases} \dfrac{C}{2A(T)}\,, & \text{if} \quad T > T_{\mathrm{c}}\,, \\[2ex] -\dfrac{C}{4A(T)}\,, & \text{if} \quad T < T_{\mathrm{c}}\,. \end{cases} \tag{14.5.27}$$

As we shall see, this specifies the natural length scale over which the spatial correlations between fluctuations decay. On both sides of the transition point ξ^2 diverges as $1/|T - T_{\mathrm{c}}|$, i.e., the correlation length itself is proportional to the inverse square root of $|T - T_{\mathrm{c}}|$.

To see the significance of ξ, we have to determine the expectation value of the spatial correlation of fluctuations,

$$\Gamma(\boldsymbol{r}) = \left\langle\delta\psi(\boldsymbol{r})\delta\psi(0)\right\rangle\,. \tag{14.5.28}$$

By recognizing that $\langle|\psi_{\boldsymbol{k}}|^2\rangle$ is the Fourier transform of this expression, an inverse transformation gives

$$\Gamma(r) \propto \int \mathrm{d}\boldsymbol{k} \frac{\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}}{1+(k\xi)^2} \propto \frac{\mathrm{e}^{-r/\xi}}{r} \,. \tag{14.5.29}$$

In the general d-dimensional case ($d \geq 2$),

$$\Gamma(r) \propto \frac{\mathrm{e}^{-r/\xi}}{r^{d-2}} \,. \tag{14.5.30}$$

Thus ξ is indeed the characteristic length of the exponential decay of correlations. Note that in the critical point, where the correlation length becomes infinitely large, the decay is no longer exponential but power-law-like.

14.5.4 Scaling Laws

Close to the critical point the behavior of the system can be characterized by a handful of parameters called critical exponents. In general, these specify the temperature and field dependence of certain thermodynamic quantities. In ferromagnetic transitions the quantities of particular interest are

$$\begin{aligned} C(T) &\propto |T-T_\mathrm{c}|^{-\alpha} \,, & M(T) &\propto (T_\mathrm{c}-T)^{\beta} \,, \\ \chi(T) &\propto |T-T_\mathrm{c}|^{-\gamma} \,, & M(H) &\propto H^{1/\delta} \,. \end{aligned} \tag{14.5.31}$$

The exponent ν that characterizes the divergence of the correlation length is defined as

$$\xi(T) \propto |T-T_\mathrm{c}|^{-\nu} \,, \tag{14.5.32}$$

while the exponent η specifying the slow spatial decay of the correlation function of the order parameter in the critical point is defined as

$$\Gamma(r) \propto \frac{1}{r^{d-2+\eta}} \,. \tag{14.5.33}$$

For those quantities that can be defined both above and below the critical temperature – such as specific heat, susceptibility, or correlation length –, the symbol used for the exponent measured in the range below T_c has an extra prime compared to the value above T_c.

According to the Landau theory, the above calculations yield $\alpha = 0$, since specific heat does not have a singularity only a finite jump, and $\beta = 1/2$, $\gamma = 1$, $\delta = 3$, $\nu = 1/2$, $\eta = 0$. However, experiments do not confirm these results. As listed in Table 14.7, the critical exponents measured in ferromagnetic materials deviate substantially from the values predicted by the Landau theory. The table contains yet another critical exponent, x, which describes the temperature dependence of the excitation energies in the spectrum of spin waves (which will be discussed in the next chapter):

Table 14.7. Experimental values for some critical exponents

	α	β	γ	δ	x
Fe	−0.12	0.38	1.33		0.37
Co		0.42	1.21		0.39
Ni	−0.10	0.38	1.32	4.2	0.39

$$D \propto (T_\mathrm{c} - T)^x \,. \tag{14.5.34}$$

In the mean-field approximation $x = 1/2$.

Using strictly valid thermodynamic relations it can be showed that the critical exponents must satisfy certain inequalities. For example, the Rushbrook (a), Griffiths (b, c), and Josephson (d, e) inequalities are[9]

$$\alpha' + 2\beta + \gamma' \geq 2 \,, \tag{14.5.35-a}$$

$$\beta(\delta + 1) \geq 2 - \alpha' \,, \tag{14.5.35-b}$$

$$\gamma' \geq \beta(\delta - 1) \,, \tag{14.5.35-c}$$

$$d\nu' \geq 2 - \alpha' \,, \tag{14.5.35-d}$$

$$d\nu \geq 2 - \alpha \,. \tag{14.5.35-e}$$

For the values obtained in the Landau theory, equalities are satisfied instead of the rigorously derived inequalities. Even though the experimental values of the critical exponents deviate from the predictions of the Landau theory, the equalities are found to hold for them, too, within experimental error.

An analysis of the experimental data also reveals another important relationship for magnetization measured as a function of the reduced temperature difference

$$t = |T - T_\mathrm{c}|/T_\mathrm{c} \tag{14.5.36}$$

and the magnetic field. Instead of studying directly the t- and H-dependence of magnetization, M/t^β can be plotted against $H/t^{\beta\delta}$. The measured values are then found to lie on a single curve in the vicinity of the critical point. Formulated mathematically,

$$M(t, H) = t^\beta f(H/t^{\beta\delta}) \,, \tag{14.5.37}$$

or alternatively

$$\frac{H}{M^\delta} = g\left(\frac{t}{M^{1/\beta}}\right) . \tag{14.5.38}$$

[9] G. S. Rushbrook, 1963, R. B. Griffiths, 1965, B. D. Josephson, 1967.

To understand this scaling property, B. Widom (1965) assumed that for small values of t and H the free energy is a generalized homogeneous function of these variables. This means that the free energy – or more precisely, its singular part coming from critical fluctuations – satisfies the equation

$$F_{\text{sing}}(\lambda^{a_t} t, \lambda^{a_H} H) = \lambda F_{\text{sing}}(t, H)\,. \tag{14.5.39}$$

This implies that specific heat, magnetization, and susceptibility show the same behavior:

$$\begin{aligned} \lambda^{2a_t} C(\lambda^{a_t} t, \lambda^{a_H} H) &= \lambda C(t, H)\,, \\ \lambda^{a_H} M(\lambda^{a_t} t, \lambda^{a_H} H) &= \lambda M(t, H)\,, \\ \lambda^{2a_H} \chi_{\text{m}}(\lambda^{a_t} t, \lambda^{a_H} H) &= \lambda \chi_{\text{m}}(t, H)\,. \end{aligned} \tag{14.5.40}$$

It is readily established that each critical exponent can be expressed with a_t and a_H:

$$\alpha = 2 - \frac{1}{a_t}\,, \quad \beta = \frac{1 - a_H}{a_t}\,, \quad \gamma = \frac{2a_H - 1}{a_t}\,, \quad \delta = \frac{a_H}{1 - a_H}\,. \tag{14.5.41}$$

This requires an appropriate choice of the scale parameter λ – for example $\lambda = t^{-1/a_t}$ for the specific heat. The above-mentioned scaling property follows directly from the equation for magnetization. It can be proved in much the same manner that the singular part of free energy also depends only on a suitably chosen combination of t and H:

$$F_{\text{sing}}(t, H) = |t|^{2-\alpha} f^{\pm}\left(\frac{H}{|t|^{\beta\delta}}\right). \tag{14.5.42}$$

Critical exponents are not independent of each other as they can be expressed with the two exponents in Widom's homogeneous function. From the equations on β and δ

$$a_t = \frac{1}{\beta}\frac{1}{\delta + 1}\,, \qquad a_H = \delta\frac{1}{\delta + 1}\,. \tag{14.5.43}$$

Substituting them into the formulas for α and γ, it is readily seen that the Rushbrook and Griffiths inequalities are replaced by equalities:

$$\alpha + 2\beta + \gamma = 2\,, \qquad \gamma = \beta(\delta - 1)\,. \tag{14.5.44}$$

In the phase transition point $t = 0$ and $H = 0$. If either t or H is nonzero, the system is no longer in the critical point, and so the correlation length becomes finite. The scaling hypothesis was understood through the insight that close to the critical point the behavior of the system is governed by the fluctuations on the only relevant scale: that of the correlation length. In 1966 L. P. Kadanoff assumed that the behavior of the system depended

only on the value of the correlation length, and that it did not make any difference whether it arose as the result of changing the magnetic field or the temperature. Within the correlation length, the length scale could be chosen at will. If instead of the natural scale of the lattice constant a its s-fold multiple is chosen as the unit, then the correlation length will be correspondingly smaller: $\xi' = \xi/s$. It is as if the scaled system were farther from the critical point – that is, if instead of the reduced temperature t and magnetic field H it was characterized by a more distant temperature t' and stronger field H'. The parameters y_t and y_H, defined by

$$t' = s^{y_t} t \qquad \text{and} \qquad H' = s^{y_H} H \tag{14.5.45}$$

are called the scaling dimensions of temperature and magnetic field. In general, the scaling dimension of any physical quantity A is determined by the relation between A', the quantity obtained by changing the length scale, and A:

$$A' = s^{y_A} A\,. \tag{14.5.46}$$

It follows from the relation $\xi \propto t^{-\nu}$ that $y_t = 1/\nu$. As for any extensive quantity, the scaling dimension for the free energy is the same as the dimension of space, therefore Kadanoff's assumption implies

$$F_{\text{sing}}(s^{y_t} t, s^{y_H} H) = s^d F_{\text{sing}}(t, H)\,. \tag{14.5.47}$$

This is just (14.5.39) – the equation asserting that free energy is a generalized homogeneous function – with the exponents

$$a_t = \frac{y_t}{d}\,, \qquad a_H = \frac{y_H}{d}\,. \tag{14.5.48}$$

Expressing a_t with the exponent of the specific heat, the Josephson inequality is also found to be replaced by an equality:

$$d\nu = 2 - \alpha\,. \tag{14.5.49}$$

In perfect analogy, it is possible to obtain a scaling formula for the correlation function, too. In addition to t and H, the distance r is also scaled now by the straightforward transformation $r' = r/s$:

$$\Gamma(s^{y_t} t, s^{y_H} H, r/s) = s^{d-2+\eta} \Gamma(t, H, r)\,. \tag{14.5.50}$$

Here η characterizes the deviation from the mean-field theory; this is why it is called *anomalous dimension*. It then follows that

$$\Gamma(r, t) = \frac{1}{r^{d-2+\eta}} g(r/\xi)\,, \tag{14.5.51}$$

or taking the Fourier transform

$$\Gamma(k, t) = \frac{1}{k^{2-\eta}} h(k\xi)\,. \tag{14.5.52}$$

The exponent η is related to the other critical exponents by the equalities

$$\gamma = (2 - \eta)\nu \tag{14.5.53}$$

and

$$2 - \eta = d\frac{\delta - 1}{\delta + 1}\,. \tag{14.5.54}$$

14.5.5 Elimination of Fluctuations and the Renormalization Group

Based on the Kadanoff theory, a physical picture was obtained for the scaling behavior of thermodynamic quantities and correlation functions close to the critical point. This allowed us to derive relations among the critical exponents. It was also understood that the deviation of experimental results from the predictions of the Landau theory are due to fluctuations. The correlation between these fluctuations becomes long-ranged in the vicinity of the critical point, and infinitely long-ranged in the critical point itself. The divergence of the correlation length gives divergent contributions to the susceptibility and other physical quantities. However, the scaling hypothesis in itself does not specify either the critical exponents or the scaling functions. It has to be complemented by K. G. Wilson's[10] formulation (1971) of a method widely used in field theory, the *renormalization-group method*. Below we shall present only the main ideas.

Since the correlation length gets macroscopically large around the critical point, fluctuations of any wavelength become important, and so the approximation that the order parameter can be determined from the minimum of the free-energy density is no longer sufficient. Characterizing the system with the effective Hamiltonian $\mathcal{H}$, we shall examine the total free energy and the partition function

$$Z = \mathrm{Tr}\exp(-\mathcal{H}/k_\mathrm{B}T)\,, \tag{14.5.55}$$

where the sum of the diagonal matrix elements is taken over all possible states of the system. Unless the critical point is at zero temperature, quantum effects are unimportant in its vicinity, so one has to deal only with thermal fluctuations that can be treated classically.[11] To this end, we shall write $\mathcal{K} = \mathcal{H}/k_\mathrm{B}T$ in a simple form in terms of the Fourier components $\psi_{\boldsymbol{k}}$ of the fluctuations of the order parameter. Starting with the form

$$\mathcal{K} = \int \mathrm{d}\boldsymbol{r}\left\{\tfrac{1}{2}r_0\psi^2(\boldsymbol{r}) + \frac{u_0}{4!}\psi^4(\boldsymbol{r}) + \tfrac{1}{2}(\boldsymbol{\nabla}\psi(\boldsymbol{r}))^2\right\}\,, \tag{14.5.56}$$

[10] Kenneth Geddes Wilson (1936–) was awarded the Nobel prize in 1982 "for his theory for critical phenomena in connection with phase transitions".

[11] For an appropriate description of phase transitions at zero temperature (called quantum phase transitions) at a quantum critical point – where the transition is not driven by temperature but by the change of the coupling constants of the Hamiltonian – quantum fluctuations must also be taken into account.

which bears some resemblance to the Landau expansion of the free-energy density (14.5.17). In terms of $\psi_{\boldsymbol{k}}$ we have

$$\mathcal{K} = \sum_{\boldsymbol{q}} \tfrac{1}{2}(r_0 + q^2)\psi_{\boldsymbol{q}}\psi_{-\boldsymbol{q}} + \frac{u_0}{4!} \sum_{\boldsymbol{q}_1,\boldsymbol{q}_2,\boldsymbol{q}_3} \psi_{\boldsymbol{q}_1}\psi_{\boldsymbol{q}_2}\psi_{\boldsymbol{q}_3}\psi_{-\boldsymbol{q}_1-\boldsymbol{q}_2-\boldsymbol{q}_3}\,. \quad (14.5.57)$$

The critical point is specified by $r_0 = 0$, while u_0 is the interaction strength of fluctuations. Both of them will be treated as phenomenological parameters. One of the most striking features of the theory is the universality of critical behavior: the critical exponents depend only on the dimensionality of the system and the order parameter, as well as the symmetries of the system – however, they are independent of the initial values of the parameters.

Taking the fluctuations into account means taking an average over fluctuations of all possible wavelengths:

$$Z = \prod_{\boldsymbol{k}} \int_{-\infty}^{\infty} \mathrm{d}\psi_{\boldsymbol{k}}\, \exp\big(-\mathcal{K}[\psi_{\boldsymbol{k}}]\big)\,. \quad (14.5.58)$$

Assuming that the most important contribution to the critical behavior comes from long-wavelength fluctuations, an average may be taken over short-wavelength (large-wave-number) fluctuations as a first step. Replacing the Brillouin zone by a sphere of radius Λ, this averaging procedure corresponds to an integration over the wavelengths between $\Lambda' = \Lambda/s$ and Λ. This way a system of fewer degrees of freedom is obtained. Then a Hamiltonian $\mathcal{H}'$ can be constructed in the smaller Brillouin zone, along with the corresponding $\mathcal{K}'$ defined as

$$\exp(-\mathcal{K}'[\psi_{\boldsymbol{k}}]) = \prod_{\Lambda/s<|\boldsymbol{k}|<\Lambda} \int_{-\infty}^{\infty} \mathrm{d}\psi_{\boldsymbol{k}}\, \exp(-\mathcal{K}[\psi_{\boldsymbol{k}}])\,. \quad (14.5.59)$$

Obviously, the partition function of this system obtained by averaging over the still allowed fluctuations must be the same as the original partition function:

$$Z_{\Lambda'}[\mathcal{K}'] = Z_{\Lambda}[\mathcal{K}]\,. \quad (14.5.60)$$

For purposes of comparing the parameters and coupling constants of the new and the original systems, the distance has to be scaled from r to $r' = r/s$ – which corresponds to scaling the wave numbers from $\boldsymbol{k}$ to $\boldsymbol{k}' = s\boldsymbol{k}$. Elimination of a part of the degrees of freedom through the above transformation is called *renormalization*, and the new coupling constant is called the *renormalized coupling*.

By continuing the procedure and eliminating further degrees of freedom – which explains the origin of the name *renormalization group*, although the group character of the transformations is not apparent in this description – the

flow of the coupling constants can be determined.[12] Things are highly simplified when the coupling constant of an interaction gets weaker and eventually vanishes as fluctuations are gradually eliminated and their effect is absorbed into the new, renormalized value of the coupling constant. Such interactions are not relevant from the viewpoint of critical behavior. However, this is generally not the case; instead the renormalized coupling tends to a finite or infinite fixed point, and often additional many-particle interactions appear as a result of renormalization. In such cases the renormalization-group transformation can be treated only numerically. The exponents that characterize the critical properties can be determined from the fixed point obtained at the end of the iteration procedure and the behavior of the system in its vicinity.

Using expressions (14.5.56) and (14.5.57) – which contain the gradient term – for the free-energy density, this renormalization program can be carried out. It turns out that in $d > 4$ dimensional systems the contribution of fluctuations is not important, and the results of the mean-field theory are recovered. In $d = 4$ dimensions the role of fluctuations is marginal, which means that the critical exponents are the same as in the mean-field approximation, however logarithmic corrections appear in the correlation functions. In the physically important case of $d < 4$ dimensions fluctuations are relevant, and so critical exponents depend on the dimensionality of space and of the order parameter. Table 14.8 shows the critical exponents determined for the three-dimensional $n = 1, 2, 3$-component Ising, XY, and isotropic Heisenberg models, compared to the values derived from the mean-field theory and the exact solution of the two-dimensional Ising model.

Table 14.8. Calculated critical exponents of the three-dimensional $n = 1, 2, 3$-component Ising, XY, and isotropic Heisenberg models, compared to the values obtained from the mean-field theory and for the two-dimensional Ising model

	α	β	γ	δ	ν	η
$d = 2$ Ising	0	1/8	7/4	15	1	1/4
$d = 3$ Ising	0.11	0.33	1.24	4.8	0.63	0.04
$d = 3$ XY	−0.01	0.35	1.32	4.8	0.67	0.04
$d = 3$ Heisenberg	−0.12	0.36	1.39	4.8	0.71	0.04
Mean-field theory	0	1/2	1	3	1/2	0

Comparison with the data listed in Table 14.7 shows that the agreement between theoretical and experimental results is fairly good for the isotropic Heisenberg model.

[12] In connection with another problem, an explicit example of the renormalization procedure will be presented in the appendix of Volume 3.

14.6 High-Temperature Expansion

On account of its initial hypotheses, the Landau theory and its improvement that takes the effects of fluctuations into consideration can describe the behavior of a magnetic system only in the vicinity of the phase transition point. We have seen that fluctuations give significant corrections to the Landau theory. Since the latter is equivalent to the mean-field theory, the question naturally arises: to what extent is the mean-field approximation justified at low and high temperatures? We shall devote the next chapter to the study of low-temperature behavior, that is the quantum mechanics of ordered magnetic structures. In the present section we shall discuss the method used in the high-temperature region.

The free energy of a system characterized by the Hamiltonian $\mathcal{H}$ can be derived from the partition function

$$Z = \operatorname{Tr} \mathrm{e}^{-\mathcal{H}/k_{\mathrm{B}}T} \tag{14.6.1}$$

as $F = -k_{\mathrm{B}}T \ln Z$, where Tr stands for the trace operator, that is, the sum of the diagonal matrix elements over a complete set of basis functions. The thermal average of a physical quantity characterized by the operator A is given by the formula

$$\langle A \rangle = \frac{\operatorname{Tr} A\mathrm{e}^{-\mathcal{H}/k_{\mathrm{B}}T}}{\operatorname{Tr} \mathrm{e}^{-\mathcal{H}/k_{\mathrm{B}}T}} \,. \tag{14.6.2}$$

At high temperatures the exponential in the partition function can be expanded into a power series of $\beta = 1/k_{\mathrm{B}}T$:

$$Z = \operatorname{Tr}(1) \left\{ 1 - \beta \langle \mathcal{H} \rangle_0 + \tfrac{1}{2}\beta^2 \langle \mathcal{H}^2 \rangle_0 + \dots \right\}, \tag{14.6.3}$$

where $\operatorname{Tr}(1)$ is the number of possible states in the complete system of functions, and $\langle B \rangle_0 = \operatorname{Tr} B / \operatorname{Tr}(1)$ stands for the thermodynamic average at infinitely large temperature. By taking the logarithm of the series, free energy can be written in terms of the so-called *cumulants*:

$$F = -k_{\mathrm{B}}T \ln \operatorname{Tr}(1) + \langle \mathcal{H} \rangle_0 - \tfrac{1}{2} \frac{1}{k_{\mathrm{B}}T} \left[\langle \mathcal{H}^2 \rangle_0 - \langle \mathcal{H} \rangle_0^2 \right] + \dots \,, \tag{14.6.4}$$

while for a thermodynamic quantity A

$$\begin{aligned} \langle A \rangle = & \left(\langle A \rangle_0 - \frac{1}{k_{\mathrm{B}}T} \langle A\mathcal{H} \rangle_0 + \tfrac{1}{2} \left(\frac{1}{k_{\mathrm{B}}T} \right)^2 \langle A\mathcal{H}^2 \rangle_0 - \dots \right) \\ & \times \left(1 + \frac{1}{k_{\mathrm{B}}T} \langle \mathcal{H} \rangle_0 + \left(\frac{1}{k_{\mathrm{B}}T} \right)^2 \left(\langle \mathcal{H} \rangle_0^2 - \tfrac{1}{2} \langle \mathcal{H}^2 \rangle_0 \right) + \dots \right). \end{aligned} \tag{14.6.5}$$

To determine the susceptibility, the contribution of the external magnetic field has to be included in the Hamiltonian, and the terms proportional to the field have to be collected. Using the same steps as in the derivation of (3.2.49), the susceptibility formula generalized to the case at hand is

$$\chi_{\alpha\beta} = V\frac{\mu_0}{k_\mathrm{B}T}\left\{\langle M_\alpha M_\beta\rangle - \langle M_\alpha\rangle\langle M_\beta\rangle\right\}, \tag{14.6.6}$$

where

$$M_\alpha = \frac{1}{V} g\mu_\mathrm{B}\sum_i S_i^\alpha \tag{14.6.7}$$

is the α component of the magnetic moment density. In the high-temperature region, where no spontaneous magnetization is present and the spin components of different directions are uncorrelated, the susceptibility tensor is diagonal:

$$\chi_\mathrm{m} = \frac{1}{V}\frac{\mu_0 g^2\mu_\mathrm{B}^2}{k_\mathrm{B}T}\Big\langle\sum_{ij} S_i^\alpha S_j^\alpha\Big\rangle. \tag{14.6.8}$$

Determining the thermal average in the previous formula for the Heisenberg model as above,

$$\chi_\mathrm{m} = \frac{N}{V}\frac{\mu_0 g^2\mu_\mathrm{B}^2 S(S+1)}{3k_\mathrm{B}T}\left[1 + \frac{2S(S+1)}{3k_\mathrm{B}T}\sum_j J_{ij} + \dots\right]. \tag{14.6.9}$$

Using the mean-field value for the Curie temperature,

$$\chi_\mathrm{m} = \frac{N}{V}\frac{\mu_0 g^2\mu_\mathrm{B}^2 S(S+1)}{3k_\mathrm{B}T}\left[1 + \frac{T_\mathrm{c}}{T} + \dots\right], \tag{14.6.10}$$

which gives the first two terms in the expansion of the Curie–Weiss susceptibility. Deviations from the mean-field theory appear in the next terms of the high-temperature expansion. By continuing the expansion to sufficiently high orders, one may try to sum up the expansion formula, that is, to fit a scaling function – in such a way that the result should be valid even in the vicinity of the transition point. The values obtained in this way for the critical temperature and the critical exponent of susceptibility in the Heisenberg model are in fair agreement with experimental data.

14.7 Magnetic Anisotropy, Domains

We have seen that the exchange interaction determines the orientation of magnetic moments only relative to each other. In crystals the orientation relative to the crystallographic axes is determined by the much weaker anisotropic terms due to relativistic spin–orbit interactions. A particularly interesting feature of ferromagnets is that in large, macroscopic samples the local magnetization does not point in the same crystallographic direction over the whole of the sample: instead, as the system starts to become ordered below the transition point, magnetization points in one of the equivalent easy axis directions around each nucleation center. The sample is thus made up of a large number of *domains* with different magnetization directions, which are separated by *domain walls*. To describe them, we shall first introduce a continuum model, and then determine the characteristic dimensions of domains and the domain-wall energy.

14.7.1 A Continuum Model of Magnetic Systems

We used a continuous order parameter in the Landau theory, and determined its magnitude from the minimum of free energy. The continuous order parameter can be kept even far from the transition point, at low temperatures – however in this region we are not concerned primarily with its magnitude (since magnetization is almost saturated here) but rather with its orientation and possible rotation.

The continuous magnetization density function is defined in terms of the magnetization density operator

$$\widehat{\boldsymbol{M}}(\boldsymbol{r}) = g\mu_{\mathrm{B}} \sum_i \boldsymbol{S}_i \delta(\boldsymbol{r} - \boldsymbol{R}_i)\,. \tag{14.7.1}$$

Even when the operator is replaced by its expectation value, the resulting expression still contains a sum of delta peaks. To obtain a coarse-grained (smoothed-out) magnetization density, the expectation value has to be averaged around $\boldsymbol{r}$ over a volume v whose radius is much larger than atomic dimensions but still small on the characteristic scale of the spatial variations of atomic magnetic moments. The expression

$$\boldsymbol{M}(\boldsymbol{r}) = \frac{1}{v} \int \langle \widehat{\boldsymbol{M}}(\boldsymbol{r}) \rangle \, \mathrm{d}\boldsymbol{r} \tag{14.7.2}$$

is indeed a classical quantity that varies slowly in space.

To express the energy with $\boldsymbol{M}(\boldsymbol{r})$, we have to start with the Heisenberg Hamiltonian. This is expressed in terms of the operator $\widehat{\boldsymbol{M}}(\boldsymbol{r})$ as

$$\mathcal{H} = -\frac{1}{(g\mu_{\mathrm{B}})^2} \int \mathrm{d}\boldsymbol{r} \int \mathrm{d}\boldsymbol{r}'\, J(\boldsymbol{r} - \boldsymbol{r}') \widehat{\boldsymbol{M}}(\boldsymbol{r}) \cdot \widehat{\boldsymbol{M}}(\boldsymbol{r}')\,, \tag{14.7.3}$$

since upon reverting to localized spins the well-known formula

$$\begin{aligned} \mathcal{H} &= -\frac{1}{(g\mu_{\mathrm{B}})^2} \int \mathrm{d}\boldsymbol{r} \int \mathrm{d}\boldsymbol{r}' \sum_{i,j} J(\boldsymbol{r} - \boldsymbol{r}')(g\mu_{\mathrm{B}})^2 \boldsymbol{S}_i \cdot \boldsymbol{S}_j \delta(\boldsymbol{r} - \boldsymbol{R}_i)\delta(\boldsymbol{r}' - \boldsymbol{R}_j) \\ &= -\sum_{i,j} J(\boldsymbol{R}_i - \boldsymbol{R}_j) \boldsymbol{S}_i \cdot \boldsymbol{S}_j \end{aligned} \tag{14.7.4}$$

is recovered. It is then plausible to assume that in terms of the density $\boldsymbol{M}(\boldsymbol{r})$ the magnetic energy can be written in the Hamiltonian-like form

$$E = -\frac{1}{(g\mu_{\mathrm{B}})^2} \int \mathrm{d}\boldsymbol{r} \int \mathrm{d}\boldsymbol{r}'\, \overline{J}(\boldsymbol{r} - \boldsymbol{r}') \boldsymbol{M}(\boldsymbol{r}) \cdot \boldsymbol{M}(\boldsymbol{r}')\,. \tag{14.7.5}$$

Just like the exchange integral $J(\boldsymbol{r})$, $\overline{J}(\boldsymbol{r})$ is also short-ranged, therefore the expansion of the position-dependent magnetization $\boldsymbol{M}(\boldsymbol{r}')$ in the integrand around $\boldsymbol{r}' = \boldsymbol{r}$ gives

$$\begin{aligned}\boldsymbol{M}(\boldsymbol{r}') = \boldsymbol{M}(\boldsymbol{r}) &+ \sum_{\mu} \frac{\partial \boldsymbol{M}(\boldsymbol{r})}{\partial r_\mu}(r'_\mu - r_\mu) \\ &+ \tfrac{1}{2}\sum_{\mu\nu} \frac{\partial^2 \boldsymbol{M}(\boldsymbol{r})}{\partial r_\mu \partial r_\nu}(r'_\mu - r_\mu)(r'_\nu - r_\nu) + \dots \,.\end{aligned} \tag{14.7.6}$$

Substituting this into the integrand, the linear term vanishes on account of the inversion symmetry of the crystal lattice. The leading term of the magnetic energy is then

$$E = \int \mathrm{d}\boldsymbol{r} \Big[- K\boldsymbol{M}(\boldsymbol{r})\boldsymbol{M}(\boldsymbol{r}) - \tfrac{1}{2}\sum_{\mu\nu} J_{\mu\nu}\boldsymbol{M}(\boldsymbol{r})\frac{\partial^2 \boldsymbol{M}(\boldsymbol{r})}{\partial r_\mu \partial r_\nu}\Big] , \tag{14.7.7}$$

where

$$K = \frac{1}{(g\mu_\mathrm{B})^2}\int \mathrm{d}\boldsymbol{r}'\, \overline{J}(\boldsymbol{r}') \,, \qquad J_{\mu\nu} = \frac{1}{(g\mu_\mathrm{B})^2}\int \mathrm{d}\boldsymbol{r}'\, \overline{J}(\boldsymbol{r}') r'_\mu r'_\nu \,. \tag{14.7.8}$$

Integration by parts gives the magnetic energy density

$$w(\boldsymbol{r}) = -K\boldsymbol{M}^2(\boldsymbol{r}) + \tfrac{1}{2}\sum_{\mu\nu} J_{\mu\nu}\frac{\partial \boldsymbol{M}(\boldsymbol{r})}{\partial r_\mu}\frac{\partial \boldsymbol{M}(\boldsymbol{r})}{\partial r_\nu} \,. \tag{14.7.9}$$

Note that this expression bears strong resemblance to the Landau expansion of free energy; however we are now interested only in the slow spatial variations in the direction of $\boldsymbol{M}(\boldsymbol{r})$ therefore the quartic term M^4 can be ignored.

The presence of $\boldsymbol{M}^2$ in the first term is the consequence of having started with the isotropic Heisenberg model. In the more general case the expression for the energy density must be invariant under the symmetry operations of the crystal. It follows from symmetry considerations that in a uniaxial crystal the combination

$$w(\boldsymbol{r}) = -K_\perp \left[M_x^2(\boldsymbol{r}) + M_y^2(\boldsymbol{r})\right] - K_\parallel M_z^2(\boldsymbol{r}) \tag{14.7.10}$$

has to appear. This determines the orientation of magnetization with respect to the crystallographic axes. The difference of the coefficients $K_\parallel$ and $K_\perp$ is therefore related to anisotropy. In cubic crystals higher-order terms have to be taken into account, since the second-order expression of cubic symmetry

$$w(\boldsymbol{r}) = -K\left[M_x^2(\boldsymbol{r}) + M_y^2(\boldsymbol{r}) + M_z^2(\boldsymbol{r})\right] \tag{14.7.11}$$

possesses full spherical symmetry, hence it does not single out a preferred direction with respect to the crystallographic axes. In the fourth order, in addition to the spherically symmetric term M^4 there are two other terms that show cubic symmetry:

$$M_x^4 + M_y^4 + M_z^4 \qquad \text{and} \qquad M_x^2M_y^2 + M_x^2M_z^2 + M_y^2M_z^2 \,. \tag{14.7.12}$$

These are not independent of each other, since a suitably chosen linear combination gives M^4. When one is not concerned with the magnitude of the moment only with its orientation, it is enough to keep either of them. Using the direction cosines α_1, α_2, α_3, the energy contribution associated with magnetic anisotropy in a cubic crystal is

$$w = K_1(\alpha_1^2\alpha_2^2 + \alpha_1^2\alpha_3^2 + \alpha_2^2\alpha_3^2) + K_2\alpha_1^2\alpha_2^2\alpha_3^2 + \dots , \tag{14.7.13}$$

which contains contributions up to the sixth order. The orientation of the moment relative to the crystallographic axes is determined by the sign of the two anisotropy constants. By minimizing the energy it is straightforward to show that for $K_1 > 0$ the magnetic moment is along one of the directions $\langle 100 \rangle$, while for $K_1 < 0$ along one of the directions $\langle 111 \rangle$.

The anisotropy constants of ferromagnetic iron and nickel measured at room temperature are given in Table 14.9. These constants depend sensitively on temperature, however even at low temperatures K_1 is positive for iron and negative for nickel. Therefore magnetization is along the edges of the cubic primitive cell in iron, and along the space diagonal in nickel.

Table 14.9. Phenomenological anisotropy constants of iron and nickel, measured at room temperature (in units of $10^3\,\mathrm{J\,m^{-3}}$)

	K_1	K_2
Fe	47.2	−0.75
Ni	−5.7	−2.3

Needless to say, the term arising from the slow spatial variations of magnetization in the energy density expression (14.7.9) also depends on the symmetries of the system. The formula obtained in the Landau expansion is recovered only in the isotropic case and for cubic crystals, where the three Cartesian coordinates are equivalent, $J_{\mu\nu} = J\delta_{\mu\nu}$, and hence the derivative term is

$$\tfrac{1}{2}J\left[\left(\frac{\partial \boldsymbol{M}(\boldsymbol{r})}{\partial x}\right)^2 + \left(\frac{\partial \boldsymbol{M}(\boldsymbol{r})}{\partial y}\right)^2 + \left(\frac{\partial \boldsymbol{M}(\boldsymbol{r})}{\partial z}\right)^2\right]. \tag{14.7.14}$$

In uniaxial crystals, where one axis is inequivalent to the two others the contribution is

$$\tfrac{1}{2}J_1\left[\left(\frac{\partial \boldsymbol{M}(\boldsymbol{r})}{\partial x}\right)^2 + \left(\frac{\partial \boldsymbol{M}(\boldsymbol{r})}{\partial y}\right)^2\right] + \tfrac{1}{2}J_2\left(\frac{\partial \boldsymbol{M}(\boldsymbol{r})}{\partial z}\right)^2. \tag{14.7.15}$$

A similar expression is found for the energy density in antiferromagnets provided the sublattice magnetization is considered as a classical vector that varies slowly in space:

$$w(\boldsymbol{r}) = K\boldsymbol{M}_{\mathrm{A}}(\boldsymbol{r}) \cdot \boldsymbol{M}_{\mathrm{B}}(\boldsymbol{r}) + \tfrac{1}{2} \sum_{\mu\nu} J_{\mu\nu} \frac{\partial \boldsymbol{M}_{\mathrm{A}}(\boldsymbol{r})}{\partial r_\mu} \frac{\partial \boldsymbol{M}_{\mathrm{B}}(\boldsymbol{r})}{\partial r_\nu} + \tfrac{1}{2} \sum_{\mu\nu} J'_{\mu\nu} \left[\frac{\partial \boldsymbol{M}_{\mathrm{A}}(\boldsymbol{r})}{\partial r_\mu} \frac{\partial \boldsymbol{M}_{\mathrm{A}}(\boldsymbol{r})}{\partial r_\nu} + \frac{\partial \boldsymbol{M}_{\mathrm{B}}(\boldsymbol{r})}{\partial r_\mu} \frac{\partial \boldsymbol{M}_{\mathrm{B}}(\boldsymbol{r})}{\partial r_\nu} \right] . \tag{14.7.16}$$

14.7.2 Magnetic Domains

In the previous calculations we ignored the energy of the magnetic field around the finite-sized magnetic sample, however this cannot be neglected in the energy balance. For homogeneous ferromagnets of finite size this field energy can be significant. As shown in Fig. 14.18, the magnetic field outside the sample – and along with it, the field energy – is reduced substantially when the sample is not homogeneously magnetized but is made up of oppositely polarized domains.

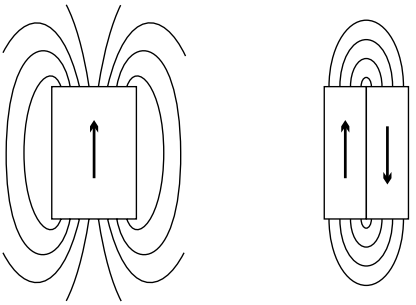


Fig. 14.18. Magnetic structures with one and two domains, and the lines of the induced magnetic field

At the boundary of the two domains there is a region in which spins are not aligned properly, therefore the formation of a wall between the domains entails an increase in magnetic energy. The latter is proportional to the surface area of the wall, while the decrease in field energy is proportional to the volume of the sample. In large samples this decrease may become dominant, in which case it is energetically more favorable to have a domain structure. However, small samples may contain a single domain.

Inside a domain the direction of magnetization is determined by anisotropy. The antiparallel orientation of the moments on the two sides of the wall is equally favorable from the viewpoint of anisotropy, however, there is a significant increase in the exchange energy. With a slow rotation over a longer distance the increase in the exchange energy can be reduced – however, this would give rise to an increase in the anisotropy energy. The competition of the two contributions determine the details of the reversal of the moment across the domain wall. Assuming uniaxial anisotropy, where the upward or

downward orientation of the moments is preferred, there are two characteristic types of domain walls. In most cases the rotation of the magnetic moment is such that it remains in the plane of the wall everywhere. Such a domain wall is called a *Bloch wall*.[13] When the rotation of the moment is in a plane perpendicular to the wall, we speak of a *Néel wall*.[14] These situations are shown in Fig. 14.19; the wall is perpendicular to the x-axis.

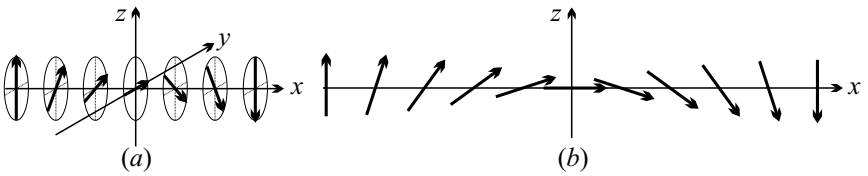


Fig. 14.19. Rotation of the magnetic moment for domain walls in the (y, z) plane: (a) Bloch wall; (b) Néel wall

In magnetically uniaxial crystals the energy density is the sum of (14.7.10) and (14.7.15). The derivative term comes from the exchange between neighboring spins, and J_1 and J_2 are related to the exchange integral. For simplicity, neglecting anisotropy in this term and the homogeneous isotropic part in the other,

$$E = \int \left\{ \frac{J}{2} \sum_{\mu} \left(\frac{\partial \boldsymbol{M}}{\partial r_\mu} \right)^2 - \frac{K}{2} M_z^2 \right\} \mathrm{d}\boldsymbol{r} \,, \tag{14.7.17}$$

where $K = K_{\parallel} - K_{\perp}$. Assuming that $K > 0$, magnetization is along the z-axis inside the domains, pointing either upward or downward. If the domain wall is in the (y, z) plane, and all the spatial variations occur in the x-direction,

$$E = \int_{-\infty}^{\infty} \left\{ \frac{J}{2} \left[\left(\frac{\partial M_x}{\partial x} \right)^2 + \left(\frac{\partial M_y}{\partial x} \right)^2 + \left(\frac{\partial M_z}{\partial x} \right)^2 \right] - \frac{K}{2} M_z^2 \right\} \mathrm{d}x \,. \tag{14.7.18}$$

Using the position-dependent polar angles θ and φ to describe the rotation of the magnetic moment,

$$M_x = M \sin\theta \cos\varphi \,, \quad M_y = M \sin\theta \sin\varphi \,, \quad M_z = M \cos\theta \tag{14.7.19}$$

in general. In terms of the variables $\theta(x)$ and $\varphi(x)$,

$$E = M^2 \int_{-\infty}^{\infty} \left\{ \frac{J}{2} \left[\left(\frac{\mathrm{d}\theta}{\mathrm{d}x} \right)^2 + \sin^2\theta \left(\frac{\mathrm{d}\varphi}{\mathrm{d}x} \right)^2 \right] - \frac{K}{2} \cos^2\theta \right\} \mathrm{d}x \,. \tag{14.7.20}$$

[13] F. Bloch, 1932.
[14] L. Néel, 1955.

Considering the spatial variations of φ first: the energy has its minimum when φ is constant. Two special cases are customarily distinguished: $\varphi = \pi/2$ corresponds to the Bloch wall, and $\varphi = 0$ to the Néel wall. In the Bloch wall the magnetic moment stays in the (y, z) plane, that is

$$M_x = 0\,, \qquad M_y = M\sin\theta\,, \qquad M_z = M\cos\theta\,. \tag{14.7.21}$$

while in the Néel wall

$$M_x = M\sin\theta\,, \qquad M_y = 0\,, \qquad M_z = M\cos\theta\,. \tag{14.7.22}$$

In both cases

$$E = M^2 \int_{-\infty}^{\infty} \left\{ \frac{J}{2}\left(\frac{\mathrm{d}\theta}{\mathrm{d}x}\right)^2 - \frac{K}{2}\cos^2\theta \right\} \mathrm{d}x\,. \tag{14.7.23}$$

The spatial variations of $\theta(x)$ have to be determined from the energy minimum. The Euler equation of the variational problem,

$$\frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial L}{\partial \theta'} - \frac{\partial L}{\partial \theta} = 0\,, \tag{14.7.24}$$

leads to the formula

$$J\frac{\mathrm{d}^2\theta}{\mathrm{d}x^2} - K\sin\theta\cos\theta = 0\,. \tag{14.7.25}$$

Integration gives

$$J\left(\frac{\mathrm{d}\theta}{\mathrm{d}x}\right)^2 - K\sin^2\theta = \text{constant}\,. \tag{14.7.26}$$

The spins are fully aligned far from the domain wall, so the following boundary condition can be imposed at infinity:

$$\begin{aligned} &\text{at } x = -\infty \qquad \theta = 0\,, \qquad \theta' = 0\,, \\ &\text{at } x = +\infty \qquad \theta = \pi\,, \qquad \theta' = 0\,. \end{aligned} \tag{14.7.27}$$

The value of the constant is therefore zero, and so

$$\left(\frac{\mathrm{d}\theta}{\mathrm{d}x}\right)^2 = \frac{K}{J}\sin^2\theta\,. \tag{14.7.28}$$

The solution of this equation that satisfies the boundary condition is

$$\cos\theta(x) = -\tanh\left(\sqrt{K/J}\,x\right)\,, \tag{14.7.29}$$

since

$$-\sin\theta\,\frac{\mathrm{d}\theta}{\mathrm{d}x} = -\sqrt{K/J}\left[1 - \tanh^2\left(\sqrt{K/J}x\right)\right] = -\sqrt{K/J}\sin^2\theta\,. \tag{14.7.30}$$

Solving (14.7.29) for $\mathrm{e}^{x/\delta}$, where $\delta = \sqrt{J/K}$,

$$\mathrm{e}^{x/\delta} = \sqrt{\frac{1-\cos\theta(x)}{1+\cos\theta(x)}} = \tan\left[\tfrac{1}{2}\theta(x)\right] , \tag{14.7.31}$$

and hence

$$\theta(x) = 2\arctan(\mathrm{e}^{x/\delta}) . \tag{14.7.32}$$

It can be immediately seen that δ is the domain-wall thickness. For obvious physical reasons, the exchange term prefers a slow rotation of the magnetic moment; therefore in itself it would lead to an infinitely thick domain wall. On the other hand, the anisotropy energy is minimal when the reversal of the moment is abrupt. The competition of these two terms leads to a finite wall thickness. Choosing values that are typical in ferromagnets for the exchange constant and the anisotropy constant, $\delta \sim 10^{-6}\,\mathrm{cm}$ is obtained. The specific thickness values for iron, cobalt, and nickel are 40, 15, and 100 nm, respectively. Therefore the number of atoms across the domain wall is on the order of 100.

The energy of the wall can also be calculated. The energy difference per unit surface area relative to the case of homogeneous magnetization along the z-axis is

$$\begin{aligned} \Delta E &= \int_{-\infty}^{\infty} \frac{M^2}{2} \left\{ K\sin^2\theta(x) - K\cos^2\theta(x) + K \right\} \mathrm{d}x \\ &= M^2 K \int_{-\infty}^{\infty} \left[1 - \tanh^2\left(\sqrt{K/J}x\right) \right] \mathrm{d}x \\ &= M^2\sqrt{JK} \int_{-\infty}^{\infty} \left[1 - \tanh^2 x \right] \mathrm{d}x = 2M^2\sqrt{JK} . \end{aligned} \tag{14.7.33}$$

This calculation gives the same energy for the Bloch wall and the Néel wall. This is because the energy contribution of magnetic dipoles has been neglected. When it is taken into account, the energy of the Bloch wall is found to be lower – except for very thin samples, since in thin magnetic films it is energetically more favorable to have the magnetization aligned with the surface of the sample everywhere, as illustrated in Fig. 14.20.

In the case of very strong uniaxial anisotropy the orientation of the moments in the interior of the domain extends all the way to the surface. In such arrangements there is always a substantial fringing field. When the anisotropy is weaker or cubic, the formation of closure domains at the surfaces is energetically more favorable, as illustrated in Fig. 14.21.

In this arrangement the magnetization component perpendicular to the surface is everywhere continuous across the boundaries of the closure domains,

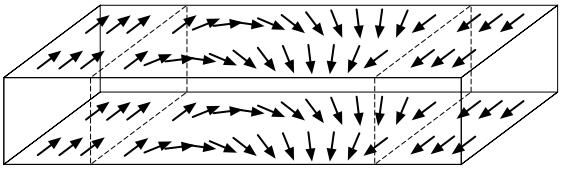


Fig. 14.20. Néel wall between two domains in a thin layer

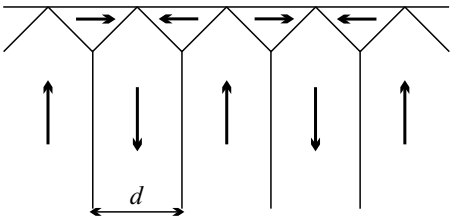


Fig. 14.21. Formation of perpendicularly magnetized closure domains close to the surface

and no fringing field appears. Using energy considerations, the dimensions of such closure domains – and through them, the width of the domains, i.e., the distance between the domain walls – can be determined. Denoting the width of the domains by d, there are L_x/d closure domains on the surface of a sample of linear size L_x. Since the volume of the closure domain is $(d^2/4)L_y$, and the anisotropy energy per unit volume is $\frac{1}{2}KM^2$, the total anisotropy energy of all closure domains is

$$E_{\text{aniso}} = \frac{L_x}{d} \cdot \frac{d^2}{4} L_y \cdot \frac{K}{2} M^2 = \frac{1}{8} K d L_x L_y M^2 \,. \tag{14.7.34}$$

Neglecting the surface energy of the closure domains in comparison with the surface energy of the L_x/d large domain walls of surface area $L_y L_z$ separated by regular distances d, and making use of the formula $2M^2\sqrt{JK}$ for the surface energy density, the wall energy is found to be

$$E_{\text{wall}} = 2M^2\sqrt{JK}\frac{L_x}{d} L_y L_z = 2M^2\sqrt{JK}\frac{L_x L_y L_z}{d} \,. \tag{14.7.35}$$

Minimizing the full energy with respect to d,

$$\frac{1}{8} K L_x L_y M^2 - 2M^2\sqrt{JK}\frac{L_x L_y L_z}{d^2} = 0 \,, \tag{14.7.36}$$

from which

$$d = 2\sqrt{L_z}\,(J/K)^{1/4} \,. \tag{14.7.37}$$

The typical value of this distance is $d \sim$ 1–10 μm, that is, the linear extension of domains corresponds to 10^4–10^5 atoms. By increasing d, further

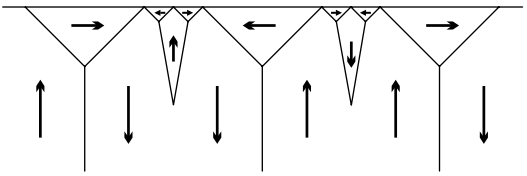


Fig. 14.22. Spike domains close to the sample surface

spike-shaped domains may appear at the surface of the sample, as shown in Fig. 14.22.

For technical magnetization curves the displacement of the domain walls and the rotation of the magnetization direction within the domains are of the utmost importance. For weak fields wall motion dominates, while for stronger fields the rotation of magnetization.

Domains can be directly observed using, for example, the powder method, in which magnetic particles trace out the pattern on the surface, or magneto-optic methods, such as those based on the Kerr effect[15] or the Faraday effect.[16] The former method is rooted in the observation that the polarization of light becomes rotated upon reflection from the surface of the sample because of the interaction with magnetic material, so the domains become visible in a polarization microscope. The rotation of the polarization plane of light is observed with methods based on the Faraday effect, too, however this time a beam penetrating through a thin magnetic layer is used. Owing to the relatively large size of domains, scanning and transmission electron microscopes may also be used for mapping the domain structure. A recent development in this field is the *magnetic force microscope* (MFM), a new version of the atomic force microscope with a magnetic stylus, designed specifically for probing magnetic materials. It has been used successfully for measuring the spatial variations of magnetization at the sample surface.

Further Reading

1. A. Aharoni, *Introduction to the Theory of Ferromagnetism*, Second Edition, Oxford Science Publications, Clarendon Press, Oxford (2001).
2. S. Chikazumi, *Physics of Ferromagnetism*, Second Edition, Oxford Science Publications, Clarendon Press, Oxford (1997).
3. A. Herpin, *Théorie du magnétisme*, Presses Universitaires de France, Paris (1968).
4. J. Jensen and A. R. Mackintosh, *Rare Earth Magnetism: Structures and Excitations*, Clarendon Press, Oxford (1991).

[15] J. Kerr, 1877.
[16] M. Faraday, 1846.

5. D. C. Jiles, *Introduction to Magnetism and Magnetic Materials*, 2nd Edition, CRC Press, Chapman & Hall, London (1998).
6. L.-P. Lévy, *Magnetism and Superconductivity*, Texts and Monographs in Physics, Springer-Verlag, Berlin (2000).
7. S. V. Vonsovskii, *Magnetism*, John Wiley & Sons, New York (1974).
8. K. Yosida, *Theory of Magnetism*, Springer Series in Solid-State Sciences, 122, Corrected 2nd Printing, Springer-Verlag, Berlin (1998).

15

Elementary Excitations in Magnetic Systems

In the previous chapter we became acquainted with magnetically ordered materials, as well as the simple description of their behavior based on the mean-field theory. The essential point was to consider only the thermal average of atomic magnetic moments and to ignore thermal and quantum fluctuations. We noted that in the vicinity of the critical point a physically as well as mathematically correct description of critical phenomena is possible only if thermal fluctuations are taken into account, and we outlined the basics of the appropriate scaling theory and of the renormalization-group transformation.

However, the mean-field theory does not provide correct quantitative results at low temperatures, either. This is because in the mean-field-theoretical treatment ordered atomic magnetic moments are assumed to point rigidly in some – quite possibly site-dependent – direction, whereas, even classically, their rotation around the effective field and the ensuing rotational degrees of freedom have to be taken into account to get a better description of the magnetic properties. This is analogous to going beyond the rigid-lattice approximation, and examining the vibrations of the ions about their equilibrium positions to understand the thermal properties of crystals. This dynamics of the spins will be examined in the present chapter. As with lattice vibrations, we shall first present a classical description of the waves formed in the system of spins, and give their quantum mechanical treatment next. We shall see that – similarly to crystalline materials, where the thermodynamic properties could be described properly with the help of the elementary excitations of bosonic character that were introduced in the quantum mechanical discussion of lattice vibrations, namely phonons – in magnetic systems, the destruction of magnetic order at finite temperature may be interpreted in terms of a gas of bosonic elementary excitations. At the end of the chapter we shall briefly discuss the anomalous behavior of low-dimensional magnetic systems, too.

15.1 Classical Spin Waves

In the mean-field-theoretical description the effective field (14.4.4) acts on the spins. It was also assumed there that the spins or magnetic moments point exactly in the direction determined by the effective field. In reality, this assumption is not justified within the classical framework, either. The fact that the thermal average of atomic moments decreases with increasing temperature should be interpreted like this: the moments turn slightly away from and precess about the direction of the internal effective magnetic field, while the component perpendicular to the field averages out to zero. To illustrate this precessional motion, we shall examine the equation of motion of atomic magnetic moments that we regard as classical vectors.

15.1.1 Ferromagnetic Spin Waves

According to classical mechanics and electrodynamics, the torque experienced by a magnetic moment $\boldsymbol{\mu}$ placed in a magnetic field $\boldsymbol{B}$ is $\boldsymbol{\mu} \times \boldsymbol{B}$, while the rate of change of the angular momentum $\boldsymbol{I}$ is given by

$$\frac{\mathrm{d}\boldsymbol{I}}{\mathrm{d}t} = \boldsymbol{\mu} \times \boldsymbol{B} \,. \tag{15.1.1}$$

We shall apply this formula to the magnetic moment $\boldsymbol{\mu}_i$ at lattice site $\boldsymbol{R}_i$ that possesses an angular momentum $\hbar \boldsymbol{S}_i$ and feels an effective field $\boldsymbol{B}_{\mathrm{eff}}$. The classical equation of motion of this spin is

$$\hbar \frac{\mathrm{d}\boldsymbol{S}_i}{\mathrm{d}t} = \boldsymbol{\mu}_i \times \boldsymbol{B}_{\mathrm{eff}} = g\mu_{\mathrm{B}}\, \boldsymbol{S}_i \times \boldsymbol{B}_{\mathrm{eff}} \,. \tag{15.1.2}$$

Using the formula (14.4.4) for the effective field, assuming zero applied field, and substituting the classical vector for the average (since the equation is classical),

$$\hbar \frac{\mathrm{d}\boldsymbol{S}_i}{\mathrm{d}t} = \boldsymbol{S}_i \times \Big(2 \sum_j J_{ij} \boldsymbol{S}_j \Big) \,. \tag{15.1.3}$$

The same equation would have emerged if we had used the quantum-mechanical equation of motion for the operator $\boldsymbol{S}_i$,

$$\frac{\mathrm{d}\boldsymbol{S}_i}{\mathrm{d}t} = \frac{\mathrm{i}}{\hbar} \left[\mathcal{H}, \boldsymbol{S}_i \right] \,, \tag{15.1.4}$$

with the Heisenberg Hamiltonian, and the commutation relations of the spin operators.

Classically, this is the equation of motion governing the precessional motion of each spin in the effective field of its neighbors. To determine its angular frequency, assume that the spins are only slightly tilted from the equilibrium value $\boldsymbol{S}_0$ common to all lattice sites,

$$\boldsymbol{S}_i = \boldsymbol{S}_0 + \delta\boldsymbol{S}_i\,, \tag{15.1.5}$$

where $\delta\boldsymbol{S}_i$ is small and perpendicular to $\boldsymbol{S}_0$. Substituting this into the equation of motion, and neglecting terms of second order in $\delta\boldsymbol{S}_i$,

$$\begin{aligned}\hbar\frac{\mathrm{d}\delta\boldsymbol{S}_i}{\mathrm{d}t} &= 2\sum_j J_{ij}\left[\delta\boldsymbol{S}_i\times\boldsymbol{S}_0 + \boldsymbol{S}_0\times\delta\boldsymbol{S}_j\right]\\ &= 2\sum_j J_{ij}\left[\delta\boldsymbol{S}_i - \delta\boldsymbol{S}_j\right]\times\boldsymbol{S}_0\,.\end{aligned} \tag{15.1.6}$$

In systems that are uniform in the ground state, this precessional motion is expected to propagate in a wave-like fashion, therefore solutions are sought in the form

$$\delta\boldsymbol{S}_i = \tfrac{1}{2}\left[\boldsymbol{A}_{\boldsymbol{k}}\mathrm{e}^{\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{R}_i-\omega_{\boldsymbol{k}}t)} + \boldsymbol{A}^*_{\boldsymbol{k}}\mathrm{e}^{-\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{R}_i-\omega_{\boldsymbol{k}}t)}\right]. \tag{15.1.7}$$

Inserting this into the equation of motion, we are led to

$$-\mathrm{i}\hbar\omega_{\boldsymbol{k}}\boldsymbol{A}_{\boldsymbol{k}} = 2\sum_j J_{ij}\left[1-\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{R}_j-\boldsymbol{R}_i)}\right]\boldsymbol{A}_{\boldsymbol{k}}\times\boldsymbol{S}_0\,, \tag{15.1.8}$$

and to a similar equation for the complex conjugate amplitude. The constraints of the vector character are satisfied by choosing the amplitude $\boldsymbol{A}_{\boldsymbol{k}}$ as

$$-\mathrm{i}\boldsymbol{A}_{\boldsymbol{k}} = \boldsymbol{A}_{\boldsymbol{k}}\times\boldsymbol{e}_0\,, \tag{15.1.9}$$

where $\boldsymbol{e}_0$ is the unit vector in the direction of the magnetization, $\boldsymbol{e}_0 = \boldsymbol{S}_0/S$, which will be chosen as the z-axis. Since the above equation asserts that $\boldsymbol{A}_{\boldsymbol{k}}$ is perpendicular to $\boldsymbol{e}_0 \equiv \hat{\boldsymbol{z}}$, $\boldsymbol{A}_{\boldsymbol{k}}$ may be written with a real $A_{\boldsymbol{k}}$ as

$$\boldsymbol{A}_{\boldsymbol{k}} = A_{\boldsymbol{k}}(\hat{\boldsymbol{x}} - \mathrm{i}\hat{\boldsymbol{y}})\,, \tag{15.1.10}$$

where $\hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{y}}$ are the unit vectors in the x and y directions, respectively. Substituting this into (15.1.7) for $\delta\boldsymbol{S}_i$, we have

$$\delta S_i^x = A_{\boldsymbol{k}}\cos(\boldsymbol{k}\cdot\boldsymbol{R}_i - \omega_{\boldsymbol{k}}t)\,, \qquad \delta S_i^y = A_{\boldsymbol{k}}\sin(\boldsymbol{k}\cdot\boldsymbol{R}_i - \omega_{\boldsymbol{k}}t)\,. \tag{15.1.11}$$

As the snapshot in Fig. 15.1 shows, the spins perform a phase-correlated precession in the plane perpendicular to the direction of the magnetization. These precessions propagating in a wave-like fashion in spin systems are called *spin waves*.

Going back to (15.1.8), the angular frequency of the precession of spins in the spin wave can be determined from

$$\hbar\omega_{\boldsymbol{k}} = 2S\sum_j J_{ij}\left[1-\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{R}_j-\boldsymbol{R}_i)}\right]. \tag{15.1.12}$$

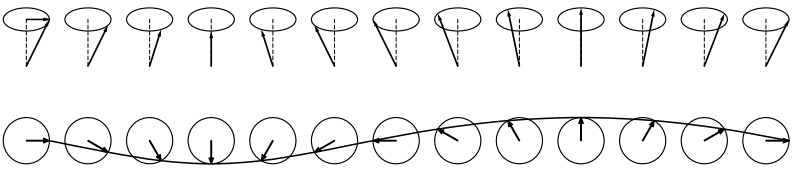


Fig. 15.1. The side and top views of the instantaneous orientation of classical spins in a spin-wave state

Expanding the exponent for large wavelengths, i.e. for small values of the wave number k, all odd terms cancel on account of the inversion symmetry of the lattice. In cubic crystals, where the three crystallographic axes are equivalent, the dispersion relation of the spin waves is isotropic in k space:

$$\hbar\omega_{\boldsymbol{k}} \sim k^2 \,. \tag{15.1.13}$$

Similarly to the classical treatment of lattice vibrations, low-frequency vibrations are obtained in the long-wavelength limit – the only difference being the quadratic, rather than linear frequency dependence of the wave number. This difference – which is a consequence of the fact that the magnetization commutes with the Hamiltonian and is therefore a conserved quantity – will play an important role in the thermodynamics of the system.

15.1.2 Spin Waves in Antiferromagnets

In antiferromagnets, just like in ferromagnets, the system of magnetic moments is expected to feature propagating waves in which the precessional motion of adjacent moments follow each other by a certain phase difference. However, because of the dissimilarity of the ground states, and since the order parameter (the sublattice magnetization) is not conserved, the dispersion relations will be essentially different.

We shall start off with the equation of motion (15.1.6) for the spin at site i, keeping in mind that in an antiferromagnet the spins are not all aligned in the same direction. We shall consider simple collinear structures in which case the antiferromagnetic order can be characterized by a wave vector $\boldsymbol{k}_0$, and the phase factor appearing in

$$\langle \boldsymbol{S}_i \rangle = \boldsymbol{S}_0 \mathrm{e}^{\mathrm{i}\boldsymbol{k}_0 \cdot \boldsymbol{R}_i} \tag{15.1.14}$$

takes the values ± 1 for the two possible spin orientations. Due to the precession of the spins, small time-dependent components perpendicular to the direction of $\boldsymbol{S}_0$ are superposed,

$$\boldsymbol{S}_i = \langle \boldsymbol{S}_i \rangle + \delta \boldsymbol{S}_i \,. \tag{15.1.15}$$

Retaining only the terms that are linear in the perpendicular component, the equation of motion reads

$$\begin{aligned}\hbar\frac{\mathrm{d}\delta\boldsymbol{S}_i}{\mathrm{d}t} &= 2\sum_j J_{ij}\left[\delta\boldsymbol{S}_i\times\langle\boldsymbol{S}_j\rangle+\langle\boldsymbol{S}_i\rangle\times\delta\boldsymbol{S}_j\right]\\ &= 2\sum_i J_{ij}\left[\delta\boldsymbol{S}_i\times\boldsymbol{S}_0\mathrm{e}^{\mathrm{i}\boldsymbol{k}_0\cdot\boldsymbol{R}_j}+\boldsymbol{S}_0\mathrm{e}^{\mathrm{i}\boldsymbol{k}_0\cdot\boldsymbol{R}_i}\times\delta\boldsymbol{S}_j\right].\end{aligned} \tag{15.1.16}$$

Assuming periodic time dependence of angular frequency ω, orienting the z-axis along the direction of $\boldsymbol{S}_0$, and introducing the variable

$$S_i^{\pm}=\delta S_i^x\pm\mathrm{i}\delta S_i^y \tag{15.1.17}$$

in the perpendicular plane, we have

$$\hbar\omega S_i^{\pm}=\mp 2\mathrm{i}S_i^{\pm}\sum_j J_{ij}S\mathrm{e}^{\mathrm{i}\boldsymbol{k}_0\cdot\boldsymbol{R}_j}\pm 2\mathrm{i}S\mathrm{e}^{\mathrm{i}\boldsymbol{k}_0\cdot\boldsymbol{R}_i}\sum_j J_{ij}S_j^{\pm}. \tag{15.1.18}$$

Again, these equations can be solved using Fourier transforms. It should be noted, however, that the components $\boldsymbol{k}$ are mixed with terms of wave vector $\boldsymbol{k}+\boldsymbol{k}_0$. The reason for this is that the magnetic cell of the antiferromagnetic structure is larger than the chemical cell and defines a Brillouin zone that is smaller than the usual one, therefore the vectors $\boldsymbol{k}$ and $\boldsymbol{k}+\boldsymbol{k}_0$, which are not equivalent in the original Brillouin zone become equivalent in the magnetic cell. For simplicity, we are considering antiferromagnets with two sublattices in which $2\boldsymbol{k}_0$ is identical with a vector in the reciprocal lattice of the chemical structure, and so further terms need not be taken into account.

Since the Brillouin zone contains $N/2$ allowed wave vectors in the two-sublattice case, we shall seek solutions of the form

$$\begin{aligned}S_i^{+} &= \sqrt{\frac{2}{N}}\sum_{\boldsymbol{k}}\left[A_{\boldsymbol{k}}\mathrm{e}^{\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{R}_i-\omega_{\boldsymbol{k}}t)}+B_{\boldsymbol{k}}\mathrm{e}^{\mathrm{i}((\boldsymbol{k}-\boldsymbol{k}_0)\cdot\boldsymbol{R}_i-\omega_{\boldsymbol{k}}t)}\right],\\ S_i^{-} &= \sqrt{\frac{2}{N}}\sum_{\boldsymbol{k}}\left[A_{\boldsymbol{k}}^{*}\mathrm{e}^{-\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{R}_i-\omega_{\boldsymbol{k}}t)}+B_{\boldsymbol{k}}^{*}\mathrm{e}^{-\mathrm{i}((\boldsymbol{k}-\boldsymbol{k}_0)\cdot\boldsymbol{R}_i-\omega_{\boldsymbol{k}}t)}\right].\end{aligned} \tag{15.1.19}$$

This corresponds to the assumption that the amplitude of the precessing component is $|A_{\boldsymbol{k}}+B_{\boldsymbol{k}}|$ on one of the sublattices, and $|A_{\boldsymbol{k}}-B_{\boldsymbol{k}}|$ on the other. Separating terms proportional to $\exp[\pm\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{R}_i-\omega_{\boldsymbol{k}}t)]$ and $\exp[\pm\mathrm{i}((\boldsymbol{k}+\boldsymbol{k}_0)\cdot\boldsymbol{R}_i-\omega_{\boldsymbol{k}}t)]$ in the equation of motion, the following relations are obtained for the coefficients:

$$\begin{aligned}\hbar\omega A_{\boldsymbol{k}} &= \pm 2SB_{\boldsymbol{k}}\sum_j J_{ij}\mathrm{e}^{\mathrm{i}\boldsymbol{k}_0\cdot(\boldsymbol{R}_j-\boldsymbol{R}_i)}\mp 2SB_{\boldsymbol{k}}\sum_j J_{ij}\mathrm{e}^{\mathrm{i}(\boldsymbol{k}-\boldsymbol{k}_0)\cdot(\boldsymbol{R}_j-\boldsymbol{R}_i)},\\ \hbar\omega B_{\boldsymbol{k}} &= \pm 2SA_{\boldsymbol{k}}\sum_j J_{ij}\mathrm{e}^{\mathrm{i}\boldsymbol{k}_0\cdot(\boldsymbol{R}_j-\boldsymbol{R}_i)}\mp 2SA_{\boldsymbol{k}}\sum_j J_{ij}\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{R}_j-\boldsymbol{R}_i)}.\end{aligned} \tag{15.1.20}$$

Making use of the Fourier transform of the exchange interaction, the precession frequency is found to be

$$\hbar\omega_{\boldsymbol{k}} = \pm 2S\sqrt{[J(\boldsymbol{k}_0) - J(\boldsymbol{k})][J(\boldsymbol{k}_0) - J(\boldsymbol{k} - \boldsymbol{k}_0)]}\,, \tag{15.1.21}$$

where

$$J(\boldsymbol{k}) = \sum_j J_{ij} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{R}_j - \boldsymbol{R}_i)}\,. \tag{15.1.22}$$

In a bipartite lattice where the nearest neighbors of each spin residing in either sublattice are located in the other sublattice, and the exchange interaction acts between nearest neighbors only,

$$J(\boldsymbol{k}_0) = z|J| \quad \text{and} \quad J(\boldsymbol{k} - \boldsymbol{k}_0) = -J(\boldsymbol{k})\,, \tag{15.1.23}$$

where z is the number of nearest neighbors. In this case the frequency takes the simple form

$$\omega_{\boldsymbol{k}} = \pm 2S\omega_0 \left(1 - \gamma_{\boldsymbol{k}}^2\right)^{1/2} \tag{15.1.24}$$

with

$$\hbar\omega_0 = \sum_j |J_{ij}| = z|J| \tag{15.1.25}$$

and

$$\gamma_{\boldsymbol{k}} = \frac{\sum_j |J_{ij}| \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{R}_i - \boldsymbol{R}_j)}}{\sum_j |J_{ij}|} = \frac{1}{z}\sum_{\boldsymbol{\delta}_j} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{\delta}_j}\,, \tag{15.1.26}$$

where $\boldsymbol{\delta}_i$ denotes the vectors pointing to the nearest neighbors. In the long-wavelength limit $\gamma_{\boldsymbol{k}}$ is very close to unity, its leading correction being of the order k^2. Thus, the dispersion relation of spin waves in antiferromagnetic materials is not quadratic but linear in k.

The snapshot of a propagating spin wave in antiferromagnets is shown in Fig. 15.2: spins precess with the same frequency but different amplitudes in the two sublattice. Because of the equivalence of the two sublattices, two types of spin waves are possible. For one, the amplitude is larger in the "up", and for the other in the "down" sublattice.

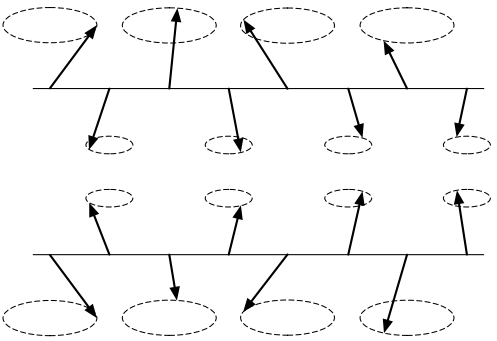


Fig. 15.2. Propagation of the two types of spin waves in a two-sublattice antiferromagnet

15.2 Quantum Mechanical Treatment of Spin Waves

Analogously to the case of phonons it may be assumed that these waves need to be quantized in the quantum mechanical treatment, in other words, that energy can change only by integral multiples of $\hbar\omega_{\boldsymbol{k}}$. Similarly to how classical vibrations lead to phonons, classical spin waves of magnetic systems will lead to *magnons*. Nevertheless, there is an essential difference: while in the quantization of lattice vibrations the canonical commutation relations for position and momentum operators implied – after the quantization of the normal coordinates – the bosonic commutation relations of phonon creation and annihilation operators, the bosonic character is only approximate for magnons, due to the special commutation relations of spin operators.

15.2.1 The Quantum Mechanics of Ferromagnetic Spin Waves

It can be easily established that the state in which the projection of each spin along a chosen direction is maximal (minimal), is an exact eigenstate of the Heisenberg Hamiltonian. To this end, consider the Hamiltonian (14.3.12) of the Heisenberg model, and choose the direction along which the spins are aligned as the quantization axis. Considering spins of magnitude S, when states with $S^z = S$ are acted upon by the part of the Hamiltonian containing the coupling between the x and y components of the spins, the result is zero, since operators S_i^+ – resulting from the transcription of S_i^x and S_j^y – cannot raise the spin projection any further. The term $S_i^z S_j^z$ just measures the state of the spins, it does not change the spin projection. This state is an eigenstate, indeed, independently of the range and sign of the exchange interaction. Thus the energy of the state with maximum spin projection is

$$E = -\sum_{i,j} J_{ij} S^2 \,. \tag{15.2.1}$$

Obviously, the state in which all spins have maximum downward projection ($S^z = -S$) is of the same energy. Denoting this state by $|0\rangle$, and the previous one by $|F\rangle$, in an isotropic system, where the total spin is conserved, further states of the same energy are easily found. Two simple examples are

$$|\psi_1\rangle = \frac{1}{\sqrt{N}} \sum_i S_i^- |F\rangle \qquad \text{and} \qquad |\psi_2\rangle = \frac{1}{\sqrt{N}} \sum_i S_i^+ |0\rangle \,. \tag{15.2.2}$$

Each expression is an eigenfunction of the square of the total spin $\boldsymbol{S}_{\rm tot} = \sum_i \boldsymbol{S}_i$, too, with the same eigenvalue $NS(NS+1)$. This shows that they all represent – with different orientations relative to the quantization axis – the same ferromagnetic state with all spins aligned.

Further states can be generated from $|\psi_1\rangle$ and $|\psi_2\rangle$ when acting upon them by the same operators that generated them. These new states belong to the same eigenvalue of $\boldsymbol{S}_{\rm tot}^2$, that is they, too, correspond to a ferromagnetic state,

just the total spin is rotated relative to the quantization axis. It can be proved by the same method that the ferromagnetic state with a spin of magnitude NS is $(2NS+1)$-fold degenerate. This has particular relevance when it is the ground state of the system. The condition for this is that exchange interactions be dominantly positive. Such a high degeneracy can, however, render the calculations difficult. This can be avoided by lifting the degeneracy by the application of an external magnetic field and breaking explicitly the SU(2) symmetry of the Heisenberg model; calculations are then performed in a finite field, and the results are eventually taken in the limit of vanishing field. Below, we shall follow the convention that the applied magnetic field is in the z direction, and so, if g is negative (which is often the case) the projection of each spin is $-S$ in the ground state. We shall denote this state by $|0\rangle$.

Excited states can be obtained from the ground state by flipping more and more spins, or partially raising their z-projections nonuniformly. Since spin flips are generated by the operators S_i^+, a general excited state is obtained by applying a series of the operators S_i^+ to the ground state. In the simplest case, a single spin, at lattice site $\boldsymbol{R}_l$ is raised from the state with projection $-S$ to that with projection $-S+1$. The resulting state is

$$|\psi_l\rangle = S_l^+|0\rangle\,. \tag{15.2.3}$$

However, this is not an eigenstate of the Heisenberg Hamiltonian. To demonstrate this, let us rewrite the Hamiltonian in terms of the spin raising and lowering operators:

$$\mathcal{H} = -\sum_{ij} J_{ij}\left[\tfrac{1}{2}\left(S_i^+S_j^- + S_i^-S_j^+\right) + S_i^zS_j^z\right] - g\mu_{\mathrm{B}}\mu_0 H\sum_i S_i^z\,. \tag{15.2.4}$$

If $j=l$, then the spin-flip terms of the type $S_i^+S_j^-$ can lower back the spin at $\boldsymbol{R}_l$, while raising another spin at a different lattice site. Therefore, only a suitable linear combination of the states ψ_l with a raised spin can be a proper eigenstate. Since the eigenstates of a Hamiltonian possessing translational symmetry may be characterized by a wave vector $\boldsymbol{k}$, we shall seek the proper eigenstate in the form

$$|\psi_{\boldsymbol{k}}\rangle = \frac{1}{\sqrt{2S}}\frac{1}{\sqrt{N}}\sum_l \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}_l}S_l^+|0\rangle\,. \tag{15.2.5}$$

When applying the Hamiltonian to this state, we have to separate those terms in which the projection of both spins connected by the exchange interaction is $-S$ from those terms in which one of these spins is raised. For the latter pairs, the $S_i^zS_j^z$ ($i=l$ or $j=l$) part of the Hamiltonian is not the only one that contributes, the $S_i^+S_l^-$ and $S_l^-S_j^+$ parts do so, too. Assuming that the system contains N spins, making use of the properties of the spin operators and the relation $J_{ij}=J_{ji}$, one obtains

$$\begin{aligned}\mathcal{H}|\psi_{\boldsymbol{k}}\rangle = {} & \frac{1}{\sqrt{2SN}} \sum_l \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}_l} \Big[- \sum_{i\neq l, j\neq l} J_{ij} S^2 - 2 \sum_{i\neq l} J_{il} S(S-1) \Big] S_l^+ |0\rangle \\ & + \frac{1}{\sqrt{2SN}} \sum_l \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}_l} \Big[- 2 \sum_{i\neq l} J_{il} S \Big] S_i^+ |0\rangle \\ & - \frac{1}{\sqrt{2SN}} \sum_l \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}_l} g\mu_\mathrm{B}\mu_0 H \big[-(N-1)S - (S-1) \big] S_l^+ |0\rangle \,. \end{aligned} \tag{15.2.6}$$

Changing the summation indices in the second term, the expression chosen for the wavefunction can be recognized on the right-hand side. Consequently, the energy eigenvalue is

$$E = E_0 + 2S \sum_j J_{ij} \left[1 - \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{R}_i - \boldsymbol{R}_j)} \right] - g\mu_\mathrm{B}\mu_0 H \,, \tag{15.2.7}$$

where

$$E_0 = - \sum_{ij} J_{ij} S^2 + g\mu_\mathrm{B}\mu_0 H N S \tag{15.2.8}$$

is the ground-state energy. The excitation energy $\hbar\omega_{\boldsymbol{k}}$ is therefore

$$\hbar\omega_{\boldsymbol{k}} = 2S \sum_j J_{ij} \left[1 - \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{R}_i - \boldsymbol{R}_j)} \right] - g\mu_\mathrm{B}\mu_0 H \,. \tag{15.2.9}$$

Aside from the second term, which is due to the magnetic field, the previous expression is analogous to (15.1.12) for the angular frequency of classical spin waves. Thus, excitation energies in the quantum mechanical treatment[1] are obtained by quantizing the classically derived frequencies in the usual way.

Considering exchange interactions of strength J between nearest neighbors only, the excitation energy is customarily written as

$$\hbar\omega_{\boldsymbol{k}} = 2JzS\,(1 - \gamma_{\boldsymbol{k}}) - g\mu_\mathrm{B}\mu_0 H \,, \tag{15.2.10}$$

where z is the number of nearest neighbors, and $\gamma_{\boldsymbol{k}}$ receives its contributions from the position vectors $\boldsymbol{\delta}_l$ of the nearest neighbors,

$$\gamma_{\boldsymbol{k}} = \frac{1}{z} \sum_{\boldsymbol{\delta}_l} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{\delta}_l} \,. \tag{15.2.11}$$

For a simple cubic lattice with lattice constant a, the factor $z(1-\gamma_{\boldsymbol{k}})$ in the excitation energy is

$$\begin{aligned} \sum_{\boldsymbol{\delta}_l} \left[1 - \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{\delta}_l} \right] &= 6 - 2\cos k_x a - 2\cos k_y a - 2\cos k_z a \\ &= 4 \left(\sin^2 \frac{k_x a}{2} + \sin^2 \frac{k_y a}{2} + \sin^2 \frac{k_z a}{2} \right) \,. \end{aligned} \tag{15.2.12}$$

[1] F. Bloch, 1930.

Expanding this for small values of the wave number, and then substituting back into the energy expression,

$$\hbar\omega_{\boldsymbol{k}} \approx 2JS\left(k_x^2 + k_y^2 + k_z^2\right)a^2 = 2JSk^2a^2 \tag{15.2.13}$$

is obtained in the absence of a magnetic field. The very same form applies to body- and face-centered cubic lattices, since for a body-centered lattice

$$\begin{aligned}\sum_l \left[1 - \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{\delta}_l}\right] &= 8 - 2\cos\tfrac{1}{2}(k_x + k_y + k_z)a - 2\cos\tfrac{1}{2}(k_x + k_y - k_z)a \\ &\quad -2\cos\tfrac{1}{2}(k_x - k_y + k_z)a - 2\cos\tfrac{1}{2}(-k_x + k_y + k_z)a \\ &= 4\Big[\sin^2\tfrac{1}{4}(k_x + k_y + k_z)a + \sin^2\tfrac{1}{4}(k_x + k_y - k_z)a \\ &\quad + \sin^2\tfrac{1}{4}(k_x - k_y + k_z)a + \sin^2\tfrac{1}{4}(-k_x + k_y + k_z)a\Big] \\ &\approx \left(k_x^2 + k_y^2 + k_z^2\right)a^2 = k^2a^2\,,\end{aligned} \tag{15.2.14}$$

while for a face-centered lattice

$$\begin{aligned}\sum_l \left[1 - \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{\delta}_l}\right] &= 12 - 4\cos\left(\tfrac{1}{2}k_xa\right)\cos\left(\tfrac{1}{2}k_ya\right) \\ &\quad - 4\cos\left(\tfrac{1}{2}k_ya\right)\cos\left(\tfrac{1}{2}k_za\right) - 4\cos\left(\tfrac{1}{2}k_za\right)\cos\left(\tfrac{1}{2}k_xa\right) \\ &\approx \left(k_x^2 + k_y^2 + k_z^2\right)a^2 = k^2a^2\,.\end{aligned} \tag{15.2.15}$$

In the long-wavelength limit, the dispersion relation of spin waves is written in the form

$$\hbar\omega_{\boldsymbol{k}} \approx Dk^2\,, \tag{15.2.16}$$

where D is the *stiffness constant* of the spin waves. The experimental values for ferromagnets of the iron group are listed in Table 15.1.

Table 15.1. The experimental values of the stiffness constant for some ferromagnets

Element	Fe	Co	Ni
D/k_B $(10^{-13}\,\mathrm{K\,cm^2})$	3.32	4.66	3.99

15.2.2 Magnons as Elementary Excitations

Upon closer examination it is noted that the wavefunction (15.2.5) of the spin-wave state is created by the operator

$$a_{\boldsymbol{k}}^{\dagger} = \frac{1}{\sqrt{2SN}} \sum_{l} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}_l} S_l^{+} \tag{15.2.17}$$

from the ground state. Taking its adjoint as the annihilation operator, their commutator is

$$\begin{aligned} \left[a_{\boldsymbol{k}}, a_{\boldsymbol{k}'}^{\dagger}\right] &= \frac{1}{2SN} \sum_{l,l'} \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}_l} \mathrm{e}^{\mathrm{i}\boldsymbol{k}'\cdot\boldsymbol{R}_{l'}} \left[S_l^{-}, S_{l'}^{+}\right] \\ &= -\frac{1}{2SN} \sum_{l} \mathrm{e}^{-\mathrm{i}(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{R}_l} 2S_l^{z} \,. \end{aligned} \tag{15.2.18}$$

At low temperatures, where the magnetization is only slightly different from the saturation value, the z component of the spins can be well approximated by $-S$, therefore

$$\left[a_{\boldsymbol{k}}, a_{\boldsymbol{k}'}^{\dagger}\right] \approx \frac{1}{N} \sum_{l} \mathrm{e}^{-\mathrm{i}(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{R}_l} = \delta_{\boldsymbol{k}\boldsymbol{k}'} \,. \tag{15.2.19}$$

Thus, in this approximation, the creation and annihilation operators of spin waves satisfy commutation relations characteristic of bosons.

Transforming the formula (15.2.17) for $a_{\boldsymbol{k}}^{\dagger}$ back into lattice representation using inverse Fourier transforms, we introduce the operators

$$a_j^{\dagger} = \frac{1}{\sqrt{2S}} S_j^{+}, \qquad a_j = \frac{1}{\sqrt{2S}} S_j^{-} \,. \tag{15.2.20}$$

In terms of these, the spin operators are

$$S_j^{+} = \sqrt{2S}\, a_j^{\dagger}, \qquad S_j^{-} = \sqrt{2S}\, a_j \,. \tag{15.2.21}$$

This pair has to be complemented by a third expression, for S_j^z. Acting on the state $|M_j\rangle$, in which the z component of the spin at site j is M_j,

$$\begin{aligned} a_j^{\dagger} a_j |M_j\rangle &= \frac{1}{2S} S_j^{+} S_j^{-} |M_j\rangle \\ &= \frac{1}{2S} \left[S(S+1) - M_j(M_j-1)\right] |M_j\rangle \\ &= \left[S + M_j + \frac{1}{2S}(S+M_j) - \frac{1}{2S}(S+M_j)^2\right] |M_j\rangle \,. \end{aligned} \tag{15.2.22}$$

If S is large enough and the quantum number M_j of the states of interest (i.e., the expectation value of S_j^z) differs little from $-S$ then the number operator $a_j^{\dagger} a_j$ measures this deviation. Thus in the space of these states

$$S_j^z = -S + a_j^{\dagger} a_j \,. \tag{15.2.23}$$

Inserting this representation of the spin operators into the Heisenberg Hamiltonian, in the present approximation

$$\begin{aligned}\mathcal{H} = &-\sum_{ij} J_{ij} S^2 + S \sum_{ij} J_{ij} \left(a_i^\dagger a_i + a_j^\dagger a_j - a_i^\dagger a_j - a_j^\dagger a_i \right) \\ &- g\mu_\mathrm{B}\mu_0 H \sum_i \left(-S + a_i^\dagger a_i \right) .\end{aligned} \tag{15.2.24}$$

Introducing the Fourier transforms of the operators $a_j^\dagger$ and a_j via

$$a_{\boldsymbol{k}}^\dagger = \frac{1}{\sqrt{N}} \sum_j \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}_j} a_j^\dagger , \qquad a_{\boldsymbol{k}} = \frac{1}{\sqrt{N}} \sum_j \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}_j} a_j , \tag{15.2.25}$$

which is equivalent to approximating the spin operator by the bosonic operator in (15.2.17), the Hamiltonian takes the simple form

$$\boxed{\mathcal{H} = E_0 + \sum_{\boldsymbol{k}} \hbar\omega_{\boldsymbol{k}} a_{\boldsymbol{k}}^\dagger a_{\boldsymbol{k}} ,} \tag{15.2.26}$$

where the expression for $\hbar\omega_{\boldsymbol{k}}$ is identical to (15.2.9) obtained for a single spin wave, and E_0 is the known energy of the ferromagnetic ground state. $\mathcal{H}$ has the same form as the Hamiltonian of a gas of bosons in which the energy of a particle with wave vector $\boldsymbol{k}$ (momentum $\hbar\boldsymbol{k}$) is $\hbar\omega_{\boldsymbol{k}}$. The elementary quanta of the latter are called *magnons*. Since the ferromagnetic ground state is exactly known, and – contrary to the case of phonons – there are no quantum fluctuations in it, the term $\frac{1}{2}$ corresponding to zero-point vibrations does not appear.

According to our previous considerations, this Hamiltonian reproduces correctly the ground state and the excited states with one raised spin. Finding further eigenstates is not so easy. Since the interaction of spins in the Heisenberg model – even in the case of uniaxial anisotropy – is such that raising the z component of one spin is always accompanied by the lowering another, thus the operator $S_\mathrm{tot}^z = \sum_i S_i^z$ commutes with the Hamiltonian,

$$[S_\mathrm{tot}^z, \mathcal{H}] = 0 , \tag{15.2.27}$$

S_tot^z is conserved. The projection of the total spin along the quantization axis can therefore be used as a quantum number to label the states. In the ground state $S_\mathrm{tot}^z = -NS$, while in one-magnon states $S_\mathrm{tot}^z = -NS + 1$, as one spin has been raised. Higher excited states are expected to have two, three, etc. raised spins. Therefore further excited states are sought in the Hilbert space with two raised spins, in the form

$$|\psi\rangle = \sum_{ij} c_{ij} S_i^+ S_j^+ |0\rangle . \tag{15.2.28}$$

If the two raised spins are not very close to each other, then these states could be considered – as suggested by the approximate form obtained for the Hamiltonian – as if two magnons were propagating in them independently.

Such states can be constructed as the products of two one-magnon states. Likewise, states containing several raised spins lend themselves to interpretation in terms of several magnons propagating independently. We shall return to the justification of this point later; now it will be assumed that an arbitrary excited state of the ferromagnet may be regarded as if it were composed of a number of independently propagating magnons. To put it otherwise: the thermodynamic behavior of ferromagnets is well approximated by that of a gas of magnons.

15.2.3 Thermodynamics of the Gas of Magnons

Since the creation of each magnon corresponds to raising a spin, in the independent spin-wave approximation the deviation of the thermodynamic average of the z component of the total spin $\boldsymbol{S}_{\mathrm{tot}}$ from the ground-state saturation value is given by the number

$$N_{\mathrm{sw}} = \sum_{\boldsymbol{k}} \langle a_{\boldsymbol{k}}^{\dagger} a_{\boldsymbol{k}} \rangle \tag{15.2.29}$$

of thermally excited magnons at temperature T, as

$$\langle S_{\mathrm{tot}}^{z} \rangle = -NS + \sum_{i} \langle a_i^{\dagger} a_i \rangle = -NS + \sum_{\boldsymbol{k}} \langle a_{\boldsymbol{k}}^{\dagger} a_{\boldsymbol{k}} \rangle \,. \tag{15.2.30}$$

Replacing summation over wave vectors by integration in the usual way, and using the Bose–Einstein distribution function for the number of magnons, one finds that the number of thermally excited magnons is

$$N_{\mathrm{sw}} = V \int \frac{\mathrm{d}\boldsymbol{k}}{(2\pi)^3} \frac{1}{\mathrm{e}^{\beta\hbar\omega_{\boldsymbol{k}}} - 1} \,, \tag{15.2.31}$$

and so the magnetization is

$$M = \frac{1}{V} g\mu_{\mathrm{B}} \langle S_{\mathrm{tot}}^{z} \rangle = \frac{N}{V} |g| \mu_{\mathrm{B}} S \left[1 - \frac{V}{NS} \int \frac{\mathrm{d}\boldsymbol{k}}{(2\pi)^3} \frac{1}{\mathrm{e}^{\beta\hbar\omega_{\boldsymbol{k}}} - 1} \right] . \tag{15.2.32}$$

In general, the temperature dependence of magnetization may be determined numerically from this formula. At low temperatures, however, one can proceed analogously to the Debye approximation used earlier for phonons. Approximating the dispersion relation by its asymptotic form in the long-wavelength limit, the integral in the square brackets may be written as

$$\frac{1}{(2\pi)^3} \int_0^{k_{\mathrm{max}}} 4\pi k^2 \, \mathrm{d}k \, \frac{1}{\mathrm{e}^{\beta D k^2} - 1} \,, \tag{15.2.33}$$

where k_{max}, the equivalent of the Debye wave number, is a cutoff related to the size of the Brillouin zone.

Unlike for phonons, the dispersion relation is now quadratic in the wave number, therefore the thermodynamics of a gas of magnons is different from that of a gas of phonons. Due to the rapid fall-off of the Bose–Einstein distribution, the cutoff can be neglected at low temperatures. Extending the upper limit of integration to infinity and introducing the new variable $\beta D k^2 = x$, the following value is found for the previous expression:

$$\begin{aligned}\frac{1}{4\pi^2}\left(\frac{k_\mathrm{B}T}{D}\right)^{3/2}\int\limits_0^\infty \frac{x^{1/2}\,\mathrm{d}x}{\mathrm{e}^x-1} &= \frac{1}{4\pi^2}\left(\frac{k_\mathrm{B}T}{D}\right)^{3/2}\frac{\sqrt{\pi}}{2}\zeta(3/2)\\ &= \left(\frac{k_\mathrm{B}T}{4\pi D}\right)^{3/2}\zeta(3/2)\,,\end{aligned} \tag{15.2.34}$$

where we have made use of (C.2.8). Inserting this expression and $D = 2JSa^2$ for the stiffness constant, as implied by (15.2.13), into (15.2.32),

$$\begin{aligned}M &= \frac{N}{V}\,|g|\mu_\mathrm{B}S\left[1-\frac{V}{NS}\zeta(3/2)\left(\frac{k_\mathrm{B}T}{4\pi D}\right)^{3/2}\right]\\ &= \frac{N}{V}\,|g|\mu_\mathrm{B}S\left[1-\frac{V}{NSa^3}\frac{\zeta(3/2)}{8\pi^{3/2}}\left(\frac{k_\mathrm{B}T}{2JS}\right)^{3/2}\right]\,.\end{aligned} \tag{15.2.35}$$

The deviation of the magnetization from its saturation value is due to thermally excited magnons. At low temperatures, the leading correction is proportional to the 3/2th power of the temperature. This result, known as the *Bloch $T^{3/2}$ law*,[2] is well confirmed by various experiments. Fitting the above function to the temperature dependence of the magnetization of nickel (shown in Fig. 14.12(a)) provides a much better agreement in the low-temperature region than the mean-field theory.

In simple cubic lattices $V = Na^3$. Inserting the numerical value $\zeta(3/2) = 2.612$,

$$M = \frac{\mu_\mathrm{B}}{a^3}\,|g|S\left[1-\frac{1}{2S}\,0.117\left(\frac{k_\mathrm{B}T}{2JS}\right)^{3/2}\right]\,. \tag{15.2.36}$$

For body-centered cubic lattices $V = Na^3/2$, and for face-centered ones $V = Na^3/4$. Introducing a multiplicative factor α that takes the values 1/2 and 1/4 in the two cases,

$$M = \frac{1}{\alpha}\,\frac{\mu_\mathrm{B}}{a^3}|g|S\left[1-\frac{\alpha}{2S}\,0.117\left(\frac{k_\mathrm{B}T}{2JS}\right)^{3/2}\right]\,. \tag{15.2.37}$$

If more realistic dispersion relations are used, corrections to the leading term appear. Assuming only nearest-neighbor exchange interactions in simple cubic lattices, the dispersion relation is

[2] F. Bloch, 1930.

$$\hbar\omega_{\boldsymbol{k}} = 4SJ\left[3 - \cos k_x a - \cos k_y a - \cos k_z a\right]. \tag{15.2.38}$$

The temperature dependence of magnetization can be determined in this case, too. In the low-temperature regime

$$M = M_0 - A_{3/2}\left(\frac{k_{\mathrm{B}}T}{2JS}\right)^{3/2} - A_{5/2}\left(\frac{k_{\mathrm{B}}T}{2JS}\right)^{5/2} - A_{7/2}\left(\frac{k_{\mathrm{B}}T}{2JS}\right)^{7/2} - \dots, \tag{15.2.39}$$

where the coefficients can be expressed in terms of Bessel functions.

To calculate the magnetic energy and the specific heat of the system, the thermal average of magnon excitation energies has to be evaluated. If the ground-state energy is neglected, the internal energy of the gas of free magnons (considered as bosonic particles) becomes

$$E = \sum_{\boldsymbol{k}} \hbar\omega_{\boldsymbol{k}} \langle a^{\dagger}_{\boldsymbol{k}} a_{\boldsymbol{k}}\rangle = \frac{V}{(2\pi)^3}\int \frac{\hbar\omega_{\boldsymbol{k}}}{\mathrm{e}^{\hbar\omega_{\boldsymbol{k}}/k_{\mathrm{B}}T} - 1}\,\mathrm{d}\boldsymbol{k}\,. \tag{15.2.40}$$

In the regime of low temperatures, this expression can be evaluated similarly to the temperature dependence of magnetization above. The result is

$$\begin{aligned} E &= \frac{V}{a^3}\frac{1}{4\pi^2}\frac{(k_{\mathrm{B}}T)^{5/2}}{(2JS)^{3/2}}\int_0^{\infty}\frac{x^{3/2}\,\mathrm{d}x}{\mathrm{e}^x - 1} \\ &= \alpha N \frac{1}{4\pi^2}\frac{(k_{\mathrm{B}}T)^{5/2}}{(2JS)^{3/2}}\frac{3\pi^{1/2}}{4}\,\zeta(5/2)\,. \end{aligned} \tag{15.2.41}$$

Due to the extra factor of $\omega_{\boldsymbol{k}}$ in (15.2.40) compared to the formula for magnetization, the energy will go with the 5/2th power of temperature. Thus the contribution of magnons to the specific heat is proportional to $T^{3/2}$:

$$C = Nk_{\mathrm{B}}A\left(\frac{k_{\mathrm{B}}T}{2JS}\right)^{3/2}, \tag{15.2.42}$$

where

$$A = \alpha\,\frac{15}{32\pi^{3/2}}\,\zeta(5/3) = 0.113\,\alpha\,. \tag{15.2.43}$$

This contribution to the low-temperature specific heat of ferromagnetic materials can indeed be easily observed in experiments, once the contribution of phonons, proportional to T^3, has been separated.

It is instructive to examine the results for the differential susceptibility of ferromagnets below the Curie temperature. To this end, we shall determine the change in the magnetization and the number of thermally excited magnons due to a magnetic field,

$$\Delta N_{\mathrm{sw}} = \sum_{\boldsymbol{k}}\left[\frac{1}{\mathrm{e}^{\beta\hbar\omega_{\boldsymbol{k}}(H)} - 1} - \frac{1}{\mathrm{e}^{\beta\hbar\omega_{\boldsymbol{k}}(H=0)} - 1}\right]. \tag{15.2.44}$$

The excitation spectrum in the presence of a magnetic field will be written as $\hbar\omega_{\boldsymbol{k}}(H) = Dk^2 + \gamma H$. Replacing summation by integration,

$$\begin{aligned}\Delta N_{\mathrm{sw}} &= V\int \frac{\mathrm{d}\boldsymbol{k}}{(2\pi)^3}\left[\frac{1}{\mathrm{e}^{\beta(Dk^2+\gamma H)}-1} - \frac{1}{\mathrm{e}^{\beta Dk^2}-1}\right] \\ &= V\int \frac{\mathrm{d}\boldsymbol{k}}{(2\pi)^3}\frac{\mathrm{e}^{\beta Dk^2}\left(1-\mathrm{e}^{\beta\gamma H}\right)}{\left(\mathrm{e}^{\beta Dk^2}-1\right)\left(\mathrm{e}^{\beta(Dk^2+\gamma H)}-1\right)}\,.\end{aligned} \tag{15.2.45}$$

For weak fields, the factor $1-\mathrm{e}^{\beta\gamma H}$ in the numerator of the integrand is proportional to H. If this were factored out, susceptibility would be proportional to the remaining integral. The latter, evaluated at $H=0$, would yield a divergent result, since close to the lower limit the integrand is proportional to $1/k^4$. Thus the term γH is also retained in the expansion valid in the vicinity of the lower limit, so we have

$$\Delta N_{\mathrm{sw}} = V\int \frac{\mathrm{d}\boldsymbol{k}}{(2\pi)^3}\frac{-\beta\gamma H}{\beta Dk^2\left(\beta Dk^2+\beta\gamma H\right)}\,. \tag{15.2.46}$$

With the new variable $x = \sqrt{D/\gamma H}\,k$,

$$\Delta N_{\mathrm{sw}} = -\frac{1}{\beta D}\left(\frac{\gamma H}{D}\right)^{1/2}\frac{V}{(2\pi)^3}\int \mathrm{d}^3x\,\frac{1}{x^2(x^2+1)}\,. \tag{15.2.47}$$

Now the integral is convergent at the lower limit. The most important point is that the magnetization is proportional to $\sqrt{H}$, hence, below T_{C}, the susceptibility exhibits a strong field dependence proportional to $1/\sqrt{H}$. In contrast to the finite value obtained from the mean-field theory, the contribution of spin waves render the susceptibility divergent in the $H\to 0$ limit. In experiments large but finite initial susceptibilities are measured instead. This is because macroscopic samples always contain domains with different directions of magnetization, and the field dependence of magnetization is in fact governed by these.

15.2.4 Rigorous Representations of Spin Operators

If the operators $a_{\boldsymbol{k}}$ and $a^{\dagger}_{\boldsymbol{k}}$ were boson operators, the operators in the lattice representation would also behave as boson operators. Obviously this cannot be so, since the z component of the spin at a given lattice site has only $(2S+1)$ allowed values. Formulas (15.2.21) and (15.2.23) can be only approximately true. If, nevertheless, the spin operators are to be represented by boson operators, more complicated expressions are needed to ensure the right commutation relations among the spin operators. Several such representations exist. Below, we shall see three commonly used forms.

In the *Holstein–Primakoff representation*[3] the spin operators are expressed as

[3] T. Holstein and H. Primakoff, 1940.

$$
\begin{aligned}
S_j^+ &= \sqrt{2S}a_j^\dagger \left(1 - a_j^\dagger a_j/2S\right)^{1/2} , \\
S_j^- &= \sqrt{2S}\left(1 - a_j^\dagger a_j/2S\right)^{1/2} a_j , \\
S_j^z &= -S + a_j^\dagger a_j .
\end{aligned}
\tag{15.2.48}
$$

With the help of the commutation relations

$$
\left[a_i, a_j^\dagger\right] = \delta_{ij} , \qquad \left[n_i, a_i\right] = -a_i , \qquad \left[n_i, a_i^\dagger\right] = a_i^\dagger \tag{15.2.49}
$$

for boson operators, it is straightforward to show that the spin operators indeed satisfy the correct commutation relations. For example, for the operators S_j^+ and S_j^-,

$$
\begin{aligned}
\left[S_j^+, S_j^-\right] &= 2S\Big[a_j^\dagger \left(1 - a_j^\dagger a_j/2S\right) a_j \\
&\qquad - \left(1 - a_j^\dagger a_j/2S\right)^{1/2} a_j a_j^\dagger \left(1 - a_j^\dagger a_j/2S\right)^{1/2}\Big] \\
&= 2S\Big[a_j^\dagger a_j \left(1 - (a_j^\dagger a_j - 1)/2S\right) \\
&\qquad - \left(1 - a_j^\dagger a_j/2S\right)^{1/2} \left(1 + a_j^\dagger a_j\right) \left(1 - a_j^\dagger a_j/2S\right)^{1/2}\Big] \\
&= 2S\left[a_j^\dagger a_j \left(1 - a_j^\dagger a_j/2S\right) + a_j^\dagger a_j/2S\right. \\
&\qquad \left. - \left(1 + a_j^\dagger a_j\right)\left(1 - a_j^\dagger a_j/2S\right)\right] \\
&= 2S\left[-1 + a_j^\dagger a_j/S\right] = 2S_j^z .
\end{aligned}
\tag{15.2.50}
$$

If we wish to go beyond the leading-order terms (which are identical to the expressions used previously), and take into account the corrections due to the factor under the square root in the Holstein–Primakoff transformation, the calculations will run into difficulties because of the square root. That is why it is often more convenient to use the *Dyson–Maleev representation*,[4] in which

$$
\begin{aligned}
S_j^+ &= \sqrt{2S}a_j^\dagger , \\
S_j^- &= \sqrt{2S}\left(1 - a_j^\dagger a_j/2S\right) a_j , \\
S_j^z &= -S + a_j^\dagger a_j .
\end{aligned}
\tag{15.2.51}
$$

Again, it is easy to demonstrate that the usual commutation relations hold among the spin operators, however, S_j^+ and S_j^- are not each other's adjoints, as they should be. Nonetheless, the Hamiltonian of the Heisenberg model proves to be Hermitian.

[4] F. J. Dyson, 1956 and S. V. Maleev, 1957.

Whether one form is used or the other, the corrections are small for large values of S, and we get back to the same expression that was used earlier. Therefore it is straightforward to assume that the gas of free magnons provides a good approximation for large spins. In reality, however, the obtained results are sufficiently precise even for $S = 1/2$ in the low-temperature regime.

In the foregoing, we gave two representations of the spin operators in terms of operators satisfying bosonic commutation relations. Further representations of the spin operators are equally possible. In the *Schwinger representation*,[5] two bosons are associated with each lattice site, and – dropping lattice indices – the vectors

$$\boldsymbol{a}^\dagger = \left(a_1^\dagger \; a_2^\dagger\right) , \qquad \boldsymbol{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \tag{15.2.52}$$

are constructed. Then the operator

$$\boldsymbol{S} = \tfrac{1}{2}\boldsymbol{a}^\dagger \cdot \boldsymbol{\sigma} \cdot \boldsymbol{a} \tag{15.2.53}$$

is defined with the help of the Pauli matrices. Its components,

$$\begin{aligned} S^z &= \tfrac{1}{2}(a_1^\dagger a_1 - a_2^\dagger a_2) , \\ S^+ &= S^x + \mathrm{i}S^y = a_1^\dagger a_2 , \\ S^- &= S^x - \mathrm{i}S^y = a_2^\dagger a_1 \end{aligned} \tag{15.2.54}$$

satisfy the commutation relations of the dimensionless angular momenta. The properties of bosons imply

$$\boldsymbol{S}^2 = S(S+1) , \tag{15.2.55}$$

where

$$S = \tfrac{1}{2}\boldsymbol{a}^\dagger \cdot \boldsymbol{a} = \tfrac{1}{2}(a_1^\dagger a_1 + a_2^\dagger a_2) . \tag{15.2.56}$$

The eigenvalues of S are integers or half-integers $(0, 1/2, 1, 3/2, 2, \dots)$. To describe the states of spin S in terms of Schwinger bosons, we have to restrict the allowed states to the subspace of the Hilbert space where the condition

$$S = \tfrac{1}{2}(a_1^\dagger a_1 + a_2^\dagger a_2) \tag{15.2.57}$$

is satisfied. If this relation is used to eliminate the boson labeled 2, and the expressions

$$\begin{aligned} a_2 &= \sqrt{2S}\left(1 - \frac{a_1^\dagger a_1}{2S}\right)^{1/2} , \\ a_2^\dagger &= \sqrt{2S}\left(1 - \frac{a_1^\dagger a_1}{2S}\right)^{1/2} \end{aligned} \tag{15.2.58}$$

[5] J. Schwinger, 1952. Julian Schwinger (1918–1994) shared the Nobel price in 1965 with S. Tomonaga and R. P. Feynman "for their fundamental work in quantum electrodynamics, with deep-ploughing consequences for the physics of elementary particles".

are substituted into (15.2.54), we obtain the Holstein–Primakoff representation of the spin operators by a single bosonic variable. If, instead, the boson labeled 1 is eliminated, then the spin operators will be expressed as

$$\begin{aligned} S_j^+ &= \sqrt{2S}\left(1 - a_j^\dagger a_j/2S\right)^{1/2} a_j\,, \\ S_j^- &= \sqrt{2S} a_j^\dagger \left(1 - a_j^\dagger a_j/2S\right)^{1/2}, \\ S_j^z &= S - a_j^\dagger a_j\,. \end{aligned} \tag{15.2.59}$$

This representation is just as good as that given in (15.2.48), but using it is practical only when the spin at the jth lattice site points upward in the ground state. We shall use this form for studying antiferromagnetic excitations.

Alternatively, in the $S = 1/2$ case spin operators can be represented by anticommuting fermion operators as

$$\boldsymbol{S}_j = \sum_{\sigma\sigma'} c_{j\sigma}^\dagger \boldsymbol{\sigma}_{\sigma\sigma'} c_{j\sigma'}\,, \tag{15.2.60}$$

or, using the explicit form of the Pauli operators

$$\begin{aligned} S_j^+ &= c_{j\uparrow}^\dagger c_{j\downarrow}, \\ S_j^- &= c_{j\downarrow}^\dagger c_{j\uparrow}, \\ S_j^z &= \tfrac{1}{2}(c_{j\uparrow}^\dagger c_{j\uparrow} - c_{j\downarrow}^\dagger c_{j\downarrow})\,. \end{aligned} \tag{15.2.61}$$

Finally, it should be mentioned that in one dimension the *Jordan–Wigner transformation*[6] can also be used to obtain a representation of spin-1/2 spin operators in terms of spinless fermions:

$$\begin{aligned} S_j^+ &= c_j^\dagger \exp\left(\mathrm{i}\pi \sum_{l<j} c_l^\dagger c_l\right), \\ S_j^- &= \exp\left(-\mathrm{i}\pi \sum_{l<j} c_l^\dagger c_l\right) c_j\,, \\ S_j^z &= c_j^\dagger c_j - \tfrac{1}{2}\,. \end{aligned} \tag{15.2.62}$$

15.2.5 Interactions Between Magnons

The expressions for spin operators in terms of magnon creation and annihilation operators are fairly complicated. As a consequence, besides bilinear terms, others containing more operators will also be included in the Heisenberg Hamiltonian, whether the Holstein–Primakoff or the Dyson–Maleev representation is used. The explicit forms of these terms are, however, different in

[6] P. Jordan and E. P. Wigner, 1928.

the two representations. Expanding the square root in the Holstein–Primakoff transformation,

$$\begin{aligned} S_j^+ &= \sqrt{2S}a_j^\dagger \left[1 - a_j^\dagger a_j/4S + \dots\right] , \\ S_j^- &= \sqrt{2S}\left[1 - a_j^\dagger a_j/4S + \dots\right] a_j , \\ S_j^z &= -S + a_j^\dagger a_j . \end{aligned} \tag{15.2.63}$$

When this is substituted into the Heisenberg Hamiltonian, and the terms quadratic and quartic in the boson operators are separated, we get

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_1 , \tag{15.2.64}$$

where $\mathcal{H}_0$ is the same as in (15.2.24),

$$\begin{aligned} \mathcal{H}_0 = &- \sum_{ij} J_{ij} S^2 + S \sum_{ij} J_{ij} \left(a_i^\dagger a_i + a_j^\dagger a_j - a_i^\dagger a_j - a_j^\dagger a_i \right) \\ &- g\mu_{\mathrm{B}}\mu_0 H \sum_i \left(-S + a_i^\dagger a_i \right) , \end{aligned} \tag{15.2.65}$$

while, making use of the relation $J_{ij} = J_{ji}$, the quartic part is

$$\mathcal{H}_1 = - \sum_{ij} J_{ij} \left(a_i^\dagger a_i a_j^\dagger a_j - \tfrac{1}{2} a_i a_j^\dagger a_j^\dagger a_j - \tfrac{1}{2} a_i^\dagger a_i a_i a_j^\dagger \right) . \tag{15.2.66}$$

As we have already seen, when expressed in terms of the Fourier transforms of the operators, $\mathcal{H}_0$ can be diagonalized. This is the Hamiltonian of the free magnon gas. On the other hand,

$$\mathcal{H}_1 = \frac{Jz}{2N} \sum_{\substack{\boldsymbol{k}_1 \boldsymbol{k}_2 \\ \boldsymbol{k}_3 \boldsymbol{k}_4}} \left(\gamma_{\boldsymbol{k}_1} + \gamma_{\boldsymbol{k}_3} - 2\gamma_{\boldsymbol{k}_1 - \boldsymbol{k}_3} \right) a_{\boldsymbol{k}_1}^\dagger a_{\boldsymbol{k}_2}^\dagger a_{\boldsymbol{k}_3} a_{\boldsymbol{k}_4} \delta(\boldsymbol{k}_1 + \boldsymbol{k}_2 - \boldsymbol{k}_3 - \boldsymbol{k}_4 + \boldsymbol{G}) \tag{15.2.67}$$

contains magnon interactions (scattering processes), so-called dynamical interactions. They have to be taken into account if the corrections to thermodynamic quantities, etc. due to magnon–magnon interactions are to be determined.

If the Dyson–Maleev transformation is applied, only such four-operator terms appear – while if the Holstein–Primakoff transformation is used instead, the quartic terms do not describe the interaction in full, the expansion of the square root brings in further terms, corresponding to multi-particle scattering.

The interactions lead to a temperature-dependent change in the magnon energy. To demonstrate this, consider the diagonal terms of the interactions that do not change the state of the magnons. For long-wavelength magnons, when umklapp processes can be neglected (only the $\boldsymbol{G} = 0$ term survives),

there are two possibilities: either $\boldsymbol{k}_1 = \boldsymbol{k}_3$ and $\boldsymbol{k}_2 = \boldsymbol{k}_4$, or $\boldsymbol{k}_1 = \boldsymbol{k}_4$ and $\boldsymbol{k}_2 = \boldsymbol{k}_3$. Rearranging the interaction part leads to

$$\mathcal{H}_1 = \frac{Jz}{N} \sum_{\boldsymbol{k}_1 \boldsymbol{k}_2} (\gamma_{\boldsymbol{k}_1} + \gamma_{\boldsymbol{k}_2} - \gamma_0 - \gamma_{\boldsymbol{k}_1 - \boldsymbol{k}_2})\, a^\dagger_{\boldsymbol{k}_1} a_{\boldsymbol{k}_1} a^\dagger_{\boldsymbol{k}_2} a_{\boldsymbol{k}_2} \,. \tag{15.2.68}$$

Assuming that the number of excited magnons fluctuates little, i.e., the terms

$$\left(a^\dagger_{\boldsymbol{k}_1} a_{\boldsymbol{k}_1} - \left\langle a^\dagger_{\boldsymbol{k}_1} a_{\boldsymbol{k}_1} \right\rangle\right) \left(a^\dagger_{\boldsymbol{k}_2} a_{\boldsymbol{k}_2} - \left\langle a^\dagger_{\boldsymbol{k}_2} a_{\boldsymbol{k}_2} \right\rangle\right) \tag{15.2.69}$$

are negligible, the diagonal part of the interaction Hamiltonian can be written in the form

$$\begin{aligned} \mathcal{H}_1 = {} & \frac{Jz}{N} \sum_{\boldsymbol{k}_1 \boldsymbol{k}_2} (\gamma_{\boldsymbol{k}_1} + \gamma_{\boldsymbol{k}_2} - \gamma_0 - \gamma_{\boldsymbol{k}_1 - \boldsymbol{k}_2})\, a^\dagger_{\boldsymbol{k}_1} a_{\boldsymbol{k}_1} \left\langle a^\dagger_{\boldsymbol{k}_2} a_{\boldsymbol{k}_2} \right\rangle \\ & + \frac{Jz}{N} \sum_{\boldsymbol{k}_1 \boldsymbol{k}_2} (\gamma_{\boldsymbol{k}_1} + \gamma_{\boldsymbol{k}_2} - \gamma_0 - \gamma_{\boldsymbol{k}_1 - \boldsymbol{k}_2}) \left\langle a^\dagger_{\boldsymbol{k}_1} a_{\boldsymbol{k}_1} \right\rangle a^\dagger_{\boldsymbol{k}_2} a_{\boldsymbol{k}_2} \\ & - \frac{Jz}{N} \sum_{\boldsymbol{k}_1 \boldsymbol{k}_2} (\gamma_{\boldsymbol{k}_1} + \gamma_{\boldsymbol{k}_2} - \gamma_0 - \gamma_{\boldsymbol{k}_1 - \boldsymbol{k}_2}) \left\langle a^\dagger_{\boldsymbol{k}_1} a_{\boldsymbol{k}_1} \right\rangle \left\langle a^\dagger_{\boldsymbol{k}_2} a_{\boldsymbol{k}_2} \right\rangle . \end{aligned} \tag{15.2.70}$$

With a change of variables the total Hamiltonian can be written as

$$\mathcal{H} = E(T) + \sum_{\boldsymbol{k}} \hbar\omega_{\boldsymbol{k}}(T) a^\dagger_{\boldsymbol{k}} a_{\boldsymbol{k}} \,, \tag{15.2.71}$$

where

$$\begin{aligned} \hbar\omega_{\boldsymbol{k}}(T) &= \hbar\omega_{\boldsymbol{k}} + \frac{2Jz}{N} \sum_{\boldsymbol{q}} (\gamma_{\boldsymbol{k}} + \gamma_{\boldsymbol{q}} - \gamma_0 - \gamma_{\boldsymbol{k}-\boldsymbol{q}}) \left\langle a^\dagger_{\boldsymbol{q}} a_{\boldsymbol{q}} \right\rangle \\ &= \hbar\omega_{\boldsymbol{k}} - \frac{1}{NS} \sum_{\boldsymbol{q}} (\hbar\omega_{\boldsymbol{k}} + \hbar\omega_{\boldsymbol{q}} - \hbar\omega_0 - \hbar\omega_{\boldsymbol{k}-\boldsymbol{q}}) \left\langle a^\dagger_{\boldsymbol{q}} a_{\boldsymbol{q}} \right\rangle . \end{aligned} \tag{15.2.72}$$

This is the energy needed to excite a magnon when other magnons are already present in the system. The pairwise interaction between the magnons renormalizes the energy of both of them, but this energy correction has to be taken into account in the total energy only once. This is taken care of by the term

$$\Delta E(T) = -\frac{Jz}{N} \sum_{\boldsymbol{k}_1 \boldsymbol{k}_2} (\gamma_{\boldsymbol{k}_1} + \gamma_{\boldsymbol{k}_2} - \gamma_0 - \gamma_{\boldsymbol{k}_1 - \boldsymbol{k}_2}) \left\langle a^\dagger_{\boldsymbol{k}_1} a_{\boldsymbol{k}_1} \right\rangle \left\langle a^\dagger_{\boldsymbol{k}_2} a_{\boldsymbol{k}_2} \right\rangle . \tag{15.2.73}$$

The renormalized energy of the magnons can be written in a simple form for a simple cubic ferromagnet with nearest-neighbor interactions. Making use of the cubic symmetry, it can be proved that

$$\hbar\omega_{\boldsymbol{k}}(T) = \hbar\omega_{\boldsymbol{k}} \left[1 - \frac{1}{2JzS^2N} \sum_{\boldsymbol{q}} \hbar\omega_{\boldsymbol{q}} \langle a^\dagger_{\boldsymbol{q}} a_{\boldsymbol{q}} \rangle \right] . \tag{15.2.74}$$

The energy of all magnons is renormalized by the same temperature-dependent factor. The correction appearing in the brackets is proportional to the thermal energy of the system due to magnetic excitations. It was shown in (15.2.41) that this quantity is proportional to $T^{5/2}$ at low temperatures. Even more generally, the correction term in the dispersion relation is proportional to k^2 in the long-wavelength regime, while its strength depends on the 5/2th power of temperature.

Over and above such processes yielding temperature-dependent energy corrections, there exist others that scatter the magnon out of its initial state. This may be interpreted as magnons having a finite lifetime. It can be shown that the inverse lifetime is proportional to k^4 in ferromagnets, that is long-wavelength magnons decay very slowly.

In equation (15.2.39) for the magnetization we saw that the corrections to the Bloch $T^{3/2}$ law involve half-integer powers of the temperature. On the other hand, the lowest-order correction due to magnon–magnon interactions is proportional to the fourth power of T (since the energy of magnons goes with $T^{5/2}$). Aside from dynamical interactions, we must take into consideration that, strictly speaking, magnons are not bosons. At each lattice site, the spin can be raised up to $2S$ times. The resulting correction is the so-called *kinematical interaction*. A rigorous treatment was presented by F. J. Dyson (1956), who showed that the first-order correction to the temperature dependence of magnetization due to kinematical interactions – similarly to that due to dynamical interactions – is proportional to T^4.

15.2.6 Two-Magnon Bound States

A simple manifestation of magnon–magnon interactions is the existence of bound states between two magnons under suitable conditions. To see how they arise, let us recall that in the ground state of the isotropic Heisenberg model with ferromagnetic ($J > 0$) exchange interaction, spins are lined up parallel, irrespective of the dimensionality of the lattice. The low-lying one-magnon excitations were obtained as linear combinations of states in which one spin is raised or lowered by one unit relative to the ground state. If spin waves were propagating independently in the lattice, then the wavefunction of the excitation of two spin waves, of wave vectors $\boldsymbol{k}$ and $\boldsymbol{k}'$, could be chosen as

$$|\psi\rangle = \frac{1}{2SN} \sum_i \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}_i} S_i^+ \sum_j \mathrm{e}^{\mathrm{i}\boldsymbol{k}'\cdot\boldsymbol{R}_j} S_j^+ |0\rangle . \tag{15.2.75}$$

This corresponds to the assumption – mentioned in connection with (15.2.28) – that the coefficients c_{ij} can be chosen in a product form. The above state is, however, not an exact eigenstate of the Hamiltonian because it contains terms

in which the two raised spins are located at adjacent sites, or, for $S > 1/2$, in which the spin at one site has been raised twice – and the result of the action of the Hamiltonian on such configurations is different from the results obtained when the two reversed spins are far apart. Since the weight of these configurations is on the order $1/N$ in the wavefunction, the excitation energy of such states is expected to be approximately

$$\Delta E = \hbar\omega_{\boldsymbol{k}} + \hbar\omega_{\boldsymbol{k}'} \,. \tag{15.2.76}$$

It can be shown that for the overwhelming majority of states with two reversed spins the energy is quite close to such a value. However, in addition to states with two nearly free spin waves, there exist other states – although with a small thermodynamic weight – that can be considered as bound states of two spin waves. For simplicity, we shall present the calculation in one dimension.

Assuming that only nearest neighbors interact, the Hamiltonian of the isotropic spin chain is written as

$$\mathcal{H} = -J \sum_l \boldsymbol{S}_l \cdot \boldsymbol{S}_{l+1} \,, \tag{15.2.77}$$

which differs from the previously used form in a factor of 2. In contrast to the foregoing, the state $|F\rangle$ in which the projection of each spin is *maximal* along the z direction will be chosen as the ferromagnetic ground state. Writing the wavefunction of the state with one lowered spin as

$$|\psi_k\rangle = \frac{1}{\sqrt{2S}} \frac{1}{\sqrt{N}} \sum_l \mathrm{e}^{\mathrm{i}kx_l} S_l^- |F\rangle \,, \tag{15.2.78}$$

the energy of one-magnon excitations is

$$\hbar\omega_k = 2SJ(1 - \cos ka) \,. \tag{15.2.79}$$

Two-magnon excitations can be obtained by reducing the z component of the total spin of the ground state by two. In state

$$|\psi_{l,l'}\rangle = S_l^- S_{l'}^- |F\rangle \tag{15.2.80}$$

the projection of the spins at sites l and l' have been reduced by one unit each. If $S = 1/2$, the operator S^- reverses the spin, and the two sites are necessarily different. If, however, $S > 1/2$, then it is possible to reduce the z component of the spin twice at the same site. These states are not eigenstates of the Heisenberg Hamiltonian, since, depending on the relative location of the two lattice sites – whether they are coincident, adjacent, or separated at a larger distance –

$$\begin{aligned}
\mathcal{H}|\psi_{l,l}\rangle &= (E_0 + 4JS)|\psi_{l,l}\rangle - J\left[S(2S-1)\right]^{1/2} \left(|\psi_{l-1,l}\rangle + |\psi_{l,l+1}\rangle\right) \,, \\
\mathcal{H}|\psi_{l,l+1}\rangle &= (E_0 + J(4S-1))|\psi_{l,l+1}\rangle - J\left[S(2S-1)\right]^{1/2} \left(|\psi_{l,l}\rangle\right. \\
&\quad \left. + |\psi_{l+1,l+1}\rangle\right) - JS\left(|\psi_{l-1,l+1}\rangle + |\psi_{l,l+2}\rangle\right) \,, \\
\mathcal{H}|\psi_{l,l'}\rangle &= (E_0 + 4JS)|\psi_{l,l'}\rangle - JS\left(|\psi_{l-1,l'}\rangle + |\psi_{l+1,l'}\rangle + |\psi_{l,l'-1}\rangle + |\psi_{l,l'+1}\rangle\right) \,.
\end{aligned} \tag{15.2.81}$$

We shall look for eigenstates expressed as linear combinations of these,

$$|\Psi\rangle = \sum_{l=1}^{N} \sum_{l'=l}^{N} c(x_l, x_{l'}) |\psi_{l,l'}\rangle \,, \qquad x_l \le x_{l'} \,. \tag{15.2.82}$$

Applying the Hamiltonian on $|\Psi\rangle$, and making use of relations (15.2.81) for the states $|\psi_{l,l'}\rangle$, equations are obtained for the coefficients $c(x_l, x_{l'})$. If the energy of the excited state is written as $E = E_0 + \hbar\omega$, we have

$$\begin{aligned}
&[\hbar\omega - 4JS] c(x_l, x_l) + J \left[S(2S-1)\right]^{1/2} \left[c(x_{l-1}, x_l) + c(x_l, x_{l+1})\right] = 0 \,, \\
&[\hbar\omega - J(4S-1)] c(x_l, x_{l+1}) + J \left[S(2S-1)\right]^{1/2} \left[c(x_l, x_l) + c(x_{l+1}, x_{l+1})\right] \\
&\qquad + JS \left[c(x_{l-1}, x_{l+1}) + c(x_l, x_{l+2})\right] = 0 \,, \\
&[\hbar\omega - 4JS] c(x_l, x_{l'}) + JS \left[c(x_{l-1}, x_{l'}) + c(x_{l+1}, x_{l'})\right. \\
&\qquad \left. + c(x_l, x_{l'-1}) + c(x_l, x_{l'+1})\right] = 0
\end{aligned} \tag{15.2.83}$$

for identical, adjacent, and further separated sites.

As a generalization of the expression $\exp(\mathrm{i}kx_l)$ used for one-magnon states and leading to plane-wave-like solutions, we shall seek the coefficient $c(x_l, x_{l'})$ in the form

$$c(x_l, x_{l'}) = A_{12} \mathrm{e}^{\mathrm{i}(k_1 x_l + k_2 x_{l'})} + A_{21} \mathrm{e}^{\mathrm{i}(k_2 x_l + k_1 x_{l'})} \,. \tag{15.2.84}$$

Both terms are necessary if we wish to obtain a symmetrized form for $x_l \le x_{l'}$. As an immediate consequence, if periodic boundary conditions are imposed, we have

$$c(x_l, x_{l'}) = c(x_{l'}, x_l + Na) \,, \tag{15.2.85}$$

or

$$A_{12} = A_{21} \mathrm{e}^{\mathrm{i}k_1 Na} = A_{21} \mathrm{e}^{-\mathrm{i}k_2 Na} \,. \tag{15.2.86}$$

The choice $A_{12} = A\mathrm{e}^{\mathrm{i}\phi/2}$, $A_{21} = A\mathrm{e}^{-\mathrm{i}\phi/2}$ leads to

$$\mathrm{e}^{\mathrm{i}k_1 Na} = \mathrm{e}^{\mathrm{i}\phi} \,, \qquad \mathrm{e}^{\mathrm{i}k_2 Na} = \mathrm{e}^{-\mathrm{i}\phi} \,, \tag{15.2.87}$$

or alternatively

$$Nak_1 = 2\pi I_1 + \phi \,, \qquad Nak_2 = 2\pi I_2 - \phi \,, \tag{15.2.88}$$

where I_1 and I_2 are integers. Physically different solutions are obtained only when I_1 and I_2 are both in the interval $[0, N)$.

Inserting this into the last equation of (15.2.83), the energy eigenvalue is found to be

$$\hbar\omega = 2JS \left[2 - (\cos k_1 a + \cos k_2 a)\right] \,. \tag{15.2.89}$$

This is apparently the same as the sum of the energies of two free magnons,

$$\hbar\omega = \hbar\omega_{k_1} + \hbar\omega_{k_2}\,. \tag{15.2.90}$$

Care must be taken however, since due to the interaction between the magnons, the two wave numbers are shifted with respect to the free-magnon values. Their sum is nonetheless a good quantum number:

$$k = k_1 + k_2 = \frac{2\pi}{Na}\,(I_1 + I_2)\,. \tag{15.2.91}$$

In terms of k and the variable $q = (k_1 - k_2)/2$, the excitation energy reads

$$\hbar\omega = 4JS\left[1 - \cos\left(\tfrac{1}{2}ka\right)\cos(qa)\right]\,. \tag{15.2.92}$$

If q is real, excitation energies are in the interval between

$$\hbar\omega_+ = 4JS\left[1 + \cos\left(\tfrac{1}{2}ka\right)\right] \quad \text{and} \quad \hbar\omega_- = 4JS\left[1 - \cos\left(\tfrac{1}{2}ka\right)\right]\,, \tag{15.2.93}$$

as shown in Fig. 15.3.

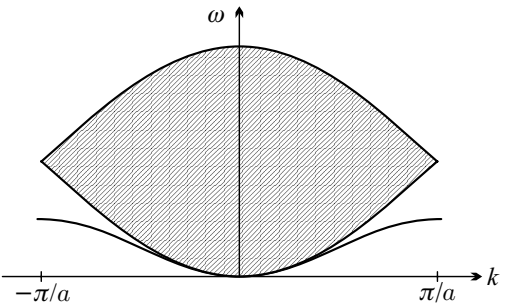


Fig. 15.3. The spectrum of excitations due to lowering two spins in a chain of $S = 1/2$ spins

The possible values of k_1 and k_2, along with the phase ϕ are determined from the condition that the first two equations of (15.2.83) (for $c(x_l, x_l)$ and $c(x_l, x_{l+1})$) should both hold. Expressing $c(x_l, x_l)$ from the first equation and substituting it into the formula for $c(x_l, x_{l+1})$, one gets

$$\begin{aligned}&[\hbar\omega - J(4S-1)]c(x_l, x_{l+1}) + JS\,(c(x_{l-1}, x_{l+1}) + c(x_l, x_{l+2}))\\ &\quad = \frac{J^2S(2S-1)}{\hbar\omega - 4JS}\,(2c(x_l, x_{l+1}) + c(x_{l-1}, x_l) + c(x_{l+1}, x_{l+2}))\,.\end{aligned} \tag{15.2.94}$$

Making use of the expression for $\hbar\omega$ and the assumption for $c(x_l, x_{l'})$, after a tedious calculation one arrives at

$$\cot(\phi/2) = \tfrac{1}{2}\left[\cot(k_1a/2) - \cot(k_2a/2)\right]\left[1 + (2S-1)\,\frac{\cos[(k_1+k_2)a/2]}{\cos[(k_1-k_2)a/2]}\right]\,. \tag{15.2.95}$$

It is readily seen that ϕ changes sign upon the interchange of k_1 and k_2, so the wavefunction is unaffected by the same interchange. For $k_1 = k_2$, $\phi = \pm\pi$, and the wave function vanishes. It is therefore sufficient to consider the case $k_1 < k_2$ ($I_1 < I_2$).

If k_2 and the corresponding I_2 are treated as known, and k_1 is varied from zero to k_2, the phase ϕ changes from 0 to π. Each wave number becomes slightly larger than the free-magnon value $2\pi I_1/Na$. The shift in energy is of order $1/N$, that is why the same regime is obtained for the continuum of such states as for free two-magnon states.

However, I_1 can only take the values $0, 1, 2, \ldots, I_2 - 2$, since the solution for $I_1 = I_2 - 1$ would be $k_1 = k_2$, implying a vanishing wavefunction, as we have already seen. Therefore, if the solutions found so far are collected for all values of I_2, we end up with $N - 1$ less states than we should. To find the missing ones, we must allow k_1 and k_2 to take complex values. We shall not go into the details of the calculation here, just convey the results: for each value of the total (resultant) wave number k there exists a complex conjugate pair $k_1, k_2 = k_1^*$ such that the corresponding excitation energy is below the bottom of the continuum of two-magnon excitations. For $S = 1/2$ the excitation energy can be given analytically as

$$\hbar\omega = \tfrac{1}{2}J\left[1 - \cos ka\right] . \tag{15.2.96}$$

In Fig. 15.3 these excitation energies are indicated by the curve below the continuum. Because of the complex wave number, these excitations correspond to states in which large amplitudes belong to configurations where the two spin flips have taken place on identical or adjacent lattice sites. These excitations can thus be regarded as bound states of two magnons. Their energy is below the continuum, because it is energetically more favorable to flip a spin that already has a flipped spin as a neighbor.

15.3 Antiferromagnetic Magnons

The description of the ground state and low-lying excited states of ferromagnetic materials was facilitated by the possibility of building the global ground state from local ground states of pairs of neighboring spins. For antiferromagnetic coupling this is no longer the case. The local ground state of two antiferromagnetically coupled spins is the singlet configuration of the spins. In this state the expectation value of the spin operator vanishes for each spin. Such local ground states cannot serve as building blocks for a global ground state with nonzero sublattice magnetization. The determination of the ground state is in fact rather difficult. Discussion of this point will be deferred to the next subsection. Here, we shall start from the classical, so-called *Néel state*, in which the z component of each spin is S in sublattice A, and $-S$ in sublattice B. In fact, this is not an eigenstate of the Heisenberg Hamiltonian, for neighboring spins cannot have maximum projections in opposite directions because

of the term

$$S_i^x S_j^x + S_i^y S_j^y = \tfrac{1}{2}\left(S_i^+ S_j^- + S_i^- S_j^+\right). \tag{15.3.1}$$

For lack of a better starting point, we shall adopt this classical ground state, and try to generalize the results obtained for ferromagnets.

15.3.1 Diagonalization of the Hamiltonian

Excited states can be studied by means of the Holstein–Primakoff or the Dyson–Maleev transformation, familiar from the treatment of ferromagnets. However, different assignments must be used on the two sublattices, with dominantly upward and downward spins ("up" and "down" sublattices). At low temperatures, the z component of the spins in sublattice A is dominantly S, while for those in sublattice B it is $-S$, hence excitations above the Néel state are created by operator S^- in sublattice A, and operator S^+ in sublattice B. Consequently, spin operators can be represented by boson operators as

$$\begin{aligned}
S_{\mathrm{A}i}^+ &= \sqrt{2S}\Big(1 - a_i^\dagger a_i/2S\Big)^{1/2} a_i\,, & S_{\mathrm{B}j}^+ &= \sqrt{2S} b_j^\dagger \Big(1 - b_j^\dagger b_j/2S\Big)^{1/2}\,,\\
S_{\mathrm{A}i}^- &= \sqrt{2S} a_i^\dagger \Big(1 - a_i^\dagger a_i/2S\Big)^{1/2}\,, & S_{\mathrm{B}j}^- &= \sqrt{2S}\Big(1 - b_j^\dagger b_j/2S\Big)^{1/2} b_j\,,\\
S_{\mathrm{A}i}^z &= S - a_i^\dagger a_i\,, & S_{\mathrm{B}j}^z &= -S + b_j^\dagger b_j\,,
\end{aligned} \tag{15.3.2}$$

or

$$\begin{aligned}
S_{\mathrm{A}i}^+ &= \sqrt{2S}\left(1 - a_i^\dagger a_i/2S\right) a_i\,, & S_{\mathrm{B}j}^+ &= \sqrt{2S} b_j^\dagger\,,\\
S_{\mathrm{A}i}^- &= \sqrt{2S} a_i^\dagger\,, & S_{\mathrm{B}j}^- &= \sqrt{2S}\left(1 - b_j^\dagger b_j/2S\right) b_j\,,\\
S_{\mathrm{A}i}^z &= S - a_i^\dagger a_i\,, & S_{\mathrm{B}j}^z &= -S + b_j^\dagger b_j\,.
\end{aligned} \tag{15.3.3}$$

Whichever choice is adopted, the correct commutation relations will be recovered for the spin operators, as long as the ladder operators a_i and b_j are bosonic in character.

We shall restrict our discussion to nearest-neighbor interactions and bipartite lattices, where the nearest neighbors of an up spin in a two-sublattice antiferromagnet are down spins in the other sublattice. Expanding the Hamiltonian through bilinear terms – that is, neglecting four-operator terms yielding magnon-magnon interactions – we have

$$\mathcal{H}_0 = 2NzJS^2 - 2JS\sum_{i\boldsymbol{\delta}}\left(a_i^\dagger a_i + b_{i+\boldsymbol{\delta}}^\dagger b_{i+\boldsymbol{\delta}} + a_i\, b_{i+\boldsymbol{\delta}} + a_i^\dagger b_{i+\boldsymbol{\delta}}^\dagger\right), \tag{15.3.4}$$

where $\boldsymbol{\delta}$ is the vector to the nearest neighbors. In terms of the Fourier transforms,

$$
\begin{aligned}
a_{\boldsymbol{k}} &= \sqrt{\frac{2}{N}} \sum_{i\in A} \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}_i} a_i \,, \qquad & a^\dagger_{\boldsymbol{k}} &= \sqrt{\frac{2}{N}} \sum_{i\in A} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}_i} a^\dagger_i \,, \\
b_{\boldsymbol{k}} &= \sqrt{\frac{2}{N}} \sum_{j\in B} \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}_j} b_j \,, \qquad & b^\dagger_{\boldsymbol{k}} &= \sqrt{\frac{2}{N}} \sum_{j\in B} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}_j} b^\dagger_j \,,
\end{aligned}
\tag{15.3.5}
$$

the Hamiltonian takes the form

$$
\mathcal{H}_0 = 2NzJS^2 - 2JzS \sum_{\boldsymbol{k}} \left[a^\dagger_{\boldsymbol{k}} a_{\boldsymbol{k}} + b^\dagger_{-\boldsymbol{k}} b_{-\boldsymbol{k}} + \gamma_{\boldsymbol{k}} \left(a_{\boldsymbol{k}} b_{-\boldsymbol{k}} + a^\dagger_{\boldsymbol{k}} b^\dagger_{-\boldsymbol{k}} \right) \right] . \tag{15.3.6}
$$

As transition to the boson operators had to be performed differently in the two sublattices, this Hamiltonian is not yet diagonal. We have to make one more transformation, the unitary Bogoliubov transformation.[7] Introducing two new creation and annihilation operators that mix the operators of the two sublattices,

$$
\begin{aligned}
\alpha_{\boldsymbol{k}} &= u_{\boldsymbol{k}} a_{\boldsymbol{k}} - v_{\boldsymbol{k}} b^\dagger_{-\boldsymbol{k}} \,, \qquad & \beta_{\boldsymbol{k}} &= u_{\boldsymbol{k}} b_{-\boldsymbol{k}} - v_{\boldsymbol{k}} a^\dagger_{\boldsymbol{k}} \,, \\
\alpha^\dagger_{\boldsymbol{k}} &= u_{\boldsymbol{k}} a^\dagger_{\boldsymbol{k}} - v_{\boldsymbol{k}} b_{-\boldsymbol{k}} \,, \qquad & \beta^\dagger_{\boldsymbol{k}} &= u_{\boldsymbol{k}} b^\dagger_{-\boldsymbol{k}} - v_{\boldsymbol{k}} a_{\boldsymbol{k}} \,.
\end{aligned}
\tag{15.3.7}
$$

The coefficients are assumed to be real; in retrospect, this will prove justified. The new operators satisfy bosonic commutation relations if $u^2_{\boldsymbol{k}} - v^2_{\boldsymbol{k}} = 1$. The inverse transformation formulas are then

$$
\begin{aligned}
a_{\boldsymbol{k}} &= u_{\boldsymbol{k}} \alpha_{\boldsymbol{k}} + v_{\boldsymbol{k}} \beta^\dagger_{\boldsymbol{k}} \,, \qquad & b_{-\boldsymbol{k}} &= u_{\boldsymbol{k}} \beta_{\boldsymbol{k}} + v_{\boldsymbol{k}} \alpha^\dagger_{\boldsymbol{k}} \,, \\
a^\dagger_{\boldsymbol{k}} &= u_{\boldsymbol{k}} \alpha^\dagger_{\boldsymbol{k}} + v_{\boldsymbol{k}} \beta_{\boldsymbol{k}}, \qquad & b^\dagger_{-\boldsymbol{k}} &= u_{\boldsymbol{k}} \beta^\dagger_{\boldsymbol{k}} + v_{\boldsymbol{k}} \alpha_{\boldsymbol{k}} \,.
\end{aligned}
\tag{15.3.8}
$$

Substituting these into (15.3.6), we find that if condition

$$
2u_{\boldsymbol{k}} v_{\boldsymbol{k}} + \gamma_{\boldsymbol{k}} \left(u^2_{\boldsymbol{k}} + v^2_{\boldsymbol{k}} \right) = 0 \tag{15.3.9}
$$

is met, only the diagonal elements in the Hamiltonian survive. Then

$$
\mathcal{H}_0 = 2NzJS^2 + 2NzJS + \sum_{\boldsymbol{k}} \hbar\omega_{\boldsymbol{k}} \left[\alpha^\dagger_{\boldsymbol{k}} \alpha_{\boldsymbol{k}} + \beta^\dagger_{\boldsymbol{k}} \beta_{\boldsymbol{k}} + 1 \right] , \tag{15.3.10}
$$

where

$$
\hbar\omega_{\boldsymbol{k}} = 2|J|zS \left[u^2_{\boldsymbol{k}} + v^2_{\boldsymbol{k}} + 2\gamma_{\boldsymbol{k}} u_{\boldsymbol{k}} v_{\boldsymbol{k}} \right] . \tag{15.3.11}
$$

The solution of (15.3.9) satisfying the auxiliary condition $u^2_{\boldsymbol{k}} - v^2_{\boldsymbol{k}} = 1$ is

$$
u^2_{\boldsymbol{k}} = \tfrac{1}{2} \left(\frac{1}{\sqrt{1-\gamma^2_{\boldsymbol{k}}}} + 1 \right), \qquad v^2_{\boldsymbol{k}} = \tfrac{1}{2} \left(\frac{1}{\sqrt{1-\gamma^2_{\boldsymbol{k}}}} - 1 \right). \tag{15.3.12}
$$

Inserting this into the excitation energy formula,

[7] N. N. Bogoliubov, 1958.

$$\hbar\omega_{\boldsymbol{k}} = 2|J|zS\left[\frac{1}{\sqrt{1-\gamma_{\boldsymbol{k}}^2}} - \frac{\gamma_{\boldsymbol{k}}^2}{\sqrt{1-\gamma_{\boldsymbol{k}}^2}}\right] = 2|J|zS\sqrt{1-\gamma_{\boldsymbol{k}}^2}\,. \tag{15.3.13}$$

In the long-wavelength limit, where $\gamma_{\boldsymbol{k}}$ is close to unity, the square root can be approximated as

$$\sqrt{1-\gamma_{\boldsymbol{k}}^2} = \sqrt{1+\gamma_{\boldsymbol{k}}}\sqrt{1-\gamma_{\boldsymbol{k}}} \approx \sqrt{2}\sqrt{1-\gamma_{\boldsymbol{k}}}\,. \tag{15.3.14}$$

The geometry-dependent factor $1-\gamma_{\boldsymbol{k}}$ was determined for cubic crystals in equations (15.2.12), (15.2.14) and (15.2.15), and in these cases

$$\sqrt{1-\gamma_{\boldsymbol{k}}^2} \approx \sqrt{\frac{2}{z}}ka\,. \tag{15.3.15}$$

Thus, as has been found in the classical limit, the dispersion relation of antiferromagnetic magnons starts linearly:

$$\hbar\omega_{\boldsymbol{k}} \approx 2JS\sqrt{2z}ka\,. \tag{15.3.16}$$

This is why antiferromagnets are strikingly different from ferromagnets, as far as thermodynamic behavior is concerned; in some aspects they are similar to a phonon gas.

In two-sublattice antiferromagnets two types of spin waves (or antiferromagnons) may therefore propagate, whose energies are equal in the absence of anisotropies and external magnetic fields. Spins in both sublattices participate in each type of excitation, however, spins in the "up" sublattice play the major role in one type, while those in the "down" sublattice in the other. If we had considered more complex antiferromagnets with more sublattices, we would have found not just two excitation branches but as many as there are sublattices. Nor will it be true any longer that each excitation branch starts at zero. Using the same terminology as for phonons, we can speak of acoustic and optical magnons.

15.3.2 The Antiferromagnetic Ground State

When expressed in terms of spin-wave creation and destruction operators, the Hamiltonian (15.3.10) of the antiferromagnetic system has, in addition to the magnon number operator, a term – "+1" in the brackets – that corresponds to the zero-point energy of the two types of magnons. In the ferromagnetic ground state, with all spins lined up parallel, there is no zero-point energy contribution. Its presence in antiferromagnets indicates that the Néel state is not the true ground state.

Taking the zero-point energy contribution into account, we have

$$E_0 = 2NzJS(S+1) + \sum_{\boldsymbol{k}} \omega_{\boldsymbol{k}} = 2NzJS\left(S + \frac{\beta}{z}\right) \tag{15.3.17}$$

for the ground-state energy, where

$$\frac{\beta}{z} = \frac{1}{N} \sum_{\boldsymbol{k}} \left(1 - \sqrt{1 - \gamma_{\boldsymbol{k}}^2} \right) . \tag{15.3.18}$$

For simple and body-centered cubic lattices $\beta \approx 0.58$. The number of nearest neighbors (the coordination number) is six in the first case and eight in the second, and so $\beta/z < 0.1$. The correction provided by zero-point vibrations is therefore less than 10% of the Néel state energy.

Due to quantum fluctuations present in the true ground state, the expectation value of the spins is slightly smaller in magnitude than the ground-state S value. To determine the zero-point spin contraction, consider the next formula for the spins in sublattice A:

$$S_{\mathrm{A}i}^z = S - a_i^\dagger a_i = S - \frac{2}{N} \sum_{\boldsymbol{k}} a_{\boldsymbol{k}}^\dagger a_{\boldsymbol{k}} . \tag{15.3.19}$$

Expressing the $a_{\boldsymbol{k}}$s in terms of the magnon creation and annihilation operators, we have

$$S_{\mathrm{A}i}^z = S - \frac{2}{N} \sum_{\boldsymbol{k}} \left(u_{\boldsymbol{k}}^2 \alpha_{\boldsymbol{k}}^\dagger \alpha_{\boldsymbol{k}} + v_{\boldsymbol{k}}^2 \beta_{\boldsymbol{k}} \beta_{\boldsymbol{k}}^\dagger + u_{\boldsymbol{k}} v_{\boldsymbol{k}} \alpha_{\boldsymbol{k}}^\dagger \beta_{-\boldsymbol{k}}^\dagger + u_{\boldsymbol{k}} v_{\boldsymbol{k}} \beta_{-\boldsymbol{k}} \alpha_{\boldsymbol{k}} \right) . \tag{15.3.20}$$

The ground-state spin reduction is the expectation value of this expression at $T = 0$. The only nonvanishing contribution comes from the term $\beta_{\boldsymbol{k}} \beta_{\boldsymbol{k}}^\dagger = 1 - \beta_{\boldsymbol{k}}^\dagger \beta_{\boldsymbol{k}}$, whence

$$\Delta S^z = S - \langle S_{\mathrm{A}i}^z \rangle = \frac{2}{N} \sum_{\boldsymbol{k}} v_{\boldsymbol{k}}^2 = \frac{1}{N} \sum_{\boldsymbol{k}} \left[\frac{1}{\sqrt{1 - \gamma_{\boldsymbol{k}}^2}} - 1 \right] . \tag{15.3.21}$$

Transforming the sum over wave vectors into an integral, and evaluating the latter numerically, 0.078 is obtained for simple cubic lattices, and 0.059 for body-centered cubic lattices. (The deviation from the mean-field theory is smaller for bcc lattices because of the larger number of nearest neighbors.) Hence, even in the worst case – that is in a cubic antiferromagnet built up of $S = 1/2$ spins – the spin is contracted by no more than 15% in the ground state. As we shall see later, this is not the case in lower-dimensional systems.

15.3.3 Antiferromagnetic Magnons at Finite Temperature

Because of the linear dispersion relation of antiferromagnetic magnons, the thermodynamic behavior of antiferromagnets is expected to differ from that of ferromagnets. Indeed, the temperature dependence of sublattice magnetization and specific heat are governed by different power laws in the two types of material.

At finite temperature, the sublattice magnetization is

$$\begin{aligned} M_{\mathrm{A}} &= \frac{1}{V} g\mu_{\mathrm{B}} \sum_i \langle S^z_{\mathrm{A}i} \rangle = \frac{1}{V}\, g\mu_{\mathrm{B}} \sum_i [S - \langle a_i^\dagger a_i \rangle] \\ &= \frac{1}{V}\, g\mu_{\mathrm{B}} \Big[\tfrac{1}{2} NS - \sum_{\boldsymbol{k}} \langle a_{\boldsymbol{k}}^\dagger a_{\boldsymbol{k}} \rangle \Big] . \end{aligned} \tag{15.3.22}$$

Changing to magnon creation and annihilation operators anew, the formula for the temperature dependence of sublattice magnetization becomes

$$M_{\mathrm{A}}(T) = M_{\mathrm{A}}(0) - g\mu_{\mathrm{B}} \frac{1}{V} \sum_{\boldsymbol{k}} \Big[u^2_{\boldsymbol{k}} \langle \alpha^\dagger_{\boldsymbol{k}} \alpha_{\boldsymbol{k}} \rangle + v^2_{\boldsymbol{k}} \langle \beta^\dagger_{\boldsymbol{k}} \beta_{\boldsymbol{k}} \rangle \Big] . \tag{15.3.23}$$

Since the two types of magnons possess the same energy, by using the Bose–Einstein statistics for the occupation numbers, and inserting the expressions for $u^2_{\boldsymbol{k}}$ and $v^2_{\boldsymbol{k}}$ from (15.3.12), we have

$$M_{\mathrm{A}}(T) = M_{\mathrm{A}}(0) - g\mu_{\mathrm{B}} \frac{1}{V} \sum_{\boldsymbol{k}} \frac{1}{\mathrm{e}^{\beta\hbar\omega_{\boldsymbol{k}}} - 1} \frac{1}{\sqrt{1 - \gamma^2_{\boldsymbol{k}}}} . \tag{15.3.24}$$

To determine the leading-order correction, the linearity of the dispersion relation is assumed at low temperatures, $\hbar\omega_{\boldsymbol{k}} = Dk$. Introduction of the new variable $x = Dk/k_{\mathrm{B}}T$ and the customary replacement of the sum by an integral leads to

$$M_{\mathrm{A}}(T) = M_{\mathrm{A}}(0) - g\mu_{\mathrm{B}} \frac{|J|zS}{D\pi^2} \left(\frac{k_{\mathrm{B}}T}{D} \right)^2 \int_0^{x_{\max}} \frac{x\,\mathrm{d}x}{\mathrm{e}^x - 1} , \tag{15.3.25}$$

where $x_{\max} = Dk_{\max}/k_{\mathrm{B}}T$, while $k_{\max}$ is determined by

$$\frac{N}{2V} = \frac{1}{(2\pi)^3} \frac{4\pi}{3}\, k^3_{\max} , \tag{15.3.26}$$

in analogy to the Debye wave number. At low temperatures, where one may extend the integration to infinity, the sublattice magnetization varies as the square of the temperature. The specific heat contribution of magnons – similarly to that of phonons, since the two dispersion relations are identical – is found to be proportional to T^3.

15.3.4 Excitations in Anisotropic Antiferromagnets

If we had determined the energy of excited states in isotropic antiferromagnets in the presence of a magnetic field, but with the constraint that in the two sublattices the spins are respectively lined up parallel and antiparallel to the applied field, we would have arrived at the simple result

$$\hbar\omega_{\boldsymbol{k}} = 2|J|zS\sqrt{1-\gamma_{\boldsymbol{k}}^2} \pm g\mu_{\mathrm{B}}B\,. \tag{15.3.27}$$

The fact that in the vicinity of $k = 0$ the excitation energies are negative in one branch signals immediately that our assumption is flawed. In isotropic models the state in which the sublattice magnetization is parallel to the magnetic field direction is not stable. As we saw in the mean-field-theoretical treatment, the magnetic moments in an isotropic antiferromagnet are not aligned with the applied magnetic field but are slightly turned with respect to the direction perpendicular to it. A uniaxial anisotropy, however, may render the alignment of the sublattice magnetization along the applied field stable. If we write the Hamiltonian as

$$\mathcal{H} = -\sum_{i,j} J_{ij}\left\{S_i^x S_j^x + S_i^y S_j^y + \Delta S_i^z S_j^z\right\}, \tag{15.3.28}$$

like in (14.3.6), the excitation spectrum can be determined applying methods similar to those used in the isotropic case. After a Holstein–Primakoff transformation from spin to boson operators, the Hamiltonian may be diagonalized by a Bogoliubov transformation. In two-sublattice antiferromagnets, if exchange takes place only between nearest neighbors located in different sublattices, the spin-wave excitation energies are found to be

$$\hbar\omega_{\boldsymbol{k}} = 2|J|zS\sqrt{\Delta^2-\gamma_{\boldsymbol{k}}^2} \pm g\mu_{\mathrm{B}}B\,. \tag{15.3.29}$$

In the absence of an applied magnetic field there is a finite gap in the magnon spectrum for $\Delta > 1$. There is no contradiction with Goldstone's theorem, for in this easy-axis anisotropic situation the ordered antiferromagnetic ground state does not break any continuous symmetry. The twofold degeneracy in the excitation spectrum is lifted by an external magnetic field. The energy in one branch decreases for increasing field intensities, and vanishes at a critical value of the magnetic field. It is at this critical field strength that the alignment of the sublattice magnetization along the easy axis of magnetization becomes unstable and that the spin-flop transition seen in the previous chapter occurs.

15.3.5 Magnons in Ferrimagnets

The magnetic excitations in ferrimagnets are examined in a bipartite lattice in the special case when the two sublattices are built up of spins of magnitude S_{A} and S_{B}, respectively, and each atom located in either sublattice is surrounded by the same z number of nearest neighbors in the other sublattice. Employing a straightforward generalization of the method used for antiferromagnets, the excitations are found to have two branches, with energies

$$\hbar\omega_{\boldsymbol{k}} = |J|z\sqrt{(S_{\mathrm{A}}-S_{\mathrm{B}})^2 + 4S_{\mathrm{A}}S_{\mathrm{B}}(1-\gamma_{\boldsymbol{k}}^2)} \pm (S_{\mathrm{A}}-S_{\mathrm{B}})\,. \tag{15.3.30}$$

For $S_A = S_B$ the spectrum of antiferromagnets is recovered. When the two spins are different, one excitation branch still starts off at zero, in accordance with Goldstone's theorem, while the energy values in the other branch are always finite. Just like for phonons, the gapless modes are called acoustic magnons, while the others are termed optical magnons.

The structure of most ferrimagnets is more complex than the one presented above. There might be more than two sublattices, with different coordination numbers and exchange integrals for each of them. In magnetite, e.g., there are six sublattices, and, accordingly, six magnon branches, of which five are optical. In the magnon spectrum of yttrium–iron garnet (YIG), 20 excitation branches are found, of which 19 are optical. In those rare-earth garnets, in which the rare-earth ions are magnetic, 32 modes are present, of which 31 are optical.

15.4 Experimental Study of Magnetic Excitations

In Chapter 13 we gave a detailed account of the experimental methods used to study lattice vibrations. We saw that by measuring the wave vector and energy of the particles (photons, neutrons etc.) in the incoming and scattered beams in scattering experiments, one can determine the spectrum of lattice vibrations created or annihilated in the scattering process. The general considerations presented there are also valid for magnetic excitations, as long as the particles in scattered beam can flip the spins of the magnetically ordered system, creating or absorbing magnetic excitations. Earlier we had also seen that neutrons, via their magnetic moments, can interact with magnetic moments localized to atoms – hence elastic neutron scattering is the most adequate method of magnetic structure determination. Therefore it comes as no surprise that inelastic neutron scattering is the method of choice for determining the dispersion relation of magnetic excitations.

According to the Van Hove formula presented in Appendix E, the double differential cross section of inelastic scattering in a vibrating lattice can be expressed in terms of the spatial and temporal correlation function of the position of atoms – or, more precisely, is proportional to its Fourier transform with respect to space and time variables. Likewise, the magnetic scattering cross section is proportional to the Fourier transform of the spatial and temporal correlation function of magnetic moments. Because of the vector character of magnetic moments, the orientation of the moments relative to the polarization direction of neutrons will be essential. We shall skip the details here – the Van Hove formula of magnetic scattering will be given in Appendix E –, and content ourselves with the result: the cross section formula contains the spatial and temporal Fourier transform of the correlation function

$$\langle S_i^\alpha(t) S_j^\beta(0) \rangle \,, \tag{15.4.1}$$

as

$$\sum_{ij} \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot(\boldsymbol{R}_i-\boldsymbol{R}_j)} \int\limits_{-\infty}^{\infty} \mathrm{d}t\, \mathrm{e}^{\mathrm{i}\varepsilon t/\hbar} \langle S_i^{\alpha}(t) S_j^{\beta}(0)\rangle \,, \tag{15.4.2}$$

where $\boldsymbol{K} = \boldsymbol{k} - \boldsymbol{k}'$ is the change of the neutron propagation vector, also known as the *scattering vector*, and $\varepsilon = E_\mathrm{f} - E_\mathrm{i}$ is the energy transfer in the process.

These correlation functions can be easily determined for systems described by the Heisenberg model if the leading term of the Holstein–Primakoff transformation is used for the spin operators. In the magnetically ordered phase the term $\langle S_i^z S_j^z\rangle$ gives rise dominantly to elastic scattering (diffraction), from which the static magnetic structure can be determined. Inelastic processes arise from the terms $\langle S_i^+ S_j^-\rangle$ and $\langle S_i^- S_j^+\rangle$. Considering, for the sake of simplicity, the ferromagnetic case, it is easily seen that, using the bosonic creation and annihilation operators, the terms appearing in the inelastic scattering cross section can be written as

$$\begin{aligned}\frac{\mathrm{d}^2\sigma}{\mathrm{d}\Omega\,\mathrm{d}\varepsilon} \propto \sum_{ij}\sum_{\boldsymbol{q}} \mathrm{e}^{-\mathrm{i}(\boldsymbol{K}+\boldsymbol{q})\cdot(\boldsymbol{R}_i-\boldsymbol{R}_j)} \int\limits_{-\infty}^{\infty} \mathrm{d}t\, \mathrm{e}^{\mathrm{i}\varepsilon t/\hbar} \langle a_{\boldsymbol{q}}^{\dagger}(t) a_{\boldsymbol{q}}(0)\rangle \\ + \sum_{ij}\sum_{\boldsymbol{q}} \mathrm{e}^{-\mathrm{i}(\boldsymbol{K}-\boldsymbol{q})\cdot(\boldsymbol{R}_i-\boldsymbol{R}_j)} \int\limits_{-\infty}^{\infty} \mathrm{d}t\, \mathrm{e}^{\mathrm{i}\varepsilon t/\hbar} \langle a_{\boldsymbol{q}}(t) a_{\boldsymbol{q}}^{\dagger}(0)\rangle \,.\end{aligned} \tag{15.4.3}$$

This expression takes a simple form when the interaction between magnons is neglected. Writing out explicitly the time dependence of the operators, which is easily obtained for free magnons, the Fourier transform of the spin–spin correlation function gives a set of Dirac delta peaks at pairs of $\boldsymbol{K}$ and ε that correspond to the creation or annihilation of magnons:

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}\Omega\,\mathrm{d}\varepsilon} \propto A\langle n_{\boldsymbol{q}}\rangle \delta(\boldsymbol{K}+\boldsymbol{q})\,\delta(\varepsilon+\hbar\omega_{\boldsymbol{q}}) + B\big(1+\langle n_{\boldsymbol{q}}\rangle\big)\delta(\boldsymbol{K}-\boldsymbol{q})\,\delta(\varepsilon-\hbar\omega_{\boldsymbol{q}})\,. \tag{15.4.4}$$

The first term – which is proportional to the magnon occupation number and has a sharp peak at $\boldsymbol{k}' = \boldsymbol{k} + \boldsymbol{q}$, $E_\mathrm{f} = E_\mathrm{i} - \hbar\omega_{\boldsymbol{q}}$ – corresponds to processes in which a magnon is absorbed by the scattered particle (neutron). The term proportional to $1 + \langle n_{\boldsymbol{q}}\rangle$ arises from processes in which a magnon is created, and therefore $\boldsymbol{k}' = \boldsymbol{k} - \boldsymbol{q}$, $E_\mathrm{f} = E_\mathrm{i} + \hbar\omega_{\boldsymbol{q}}$.

By measuring the peaks as a function of the transferred energy and the scattering angle, the dispersion relation for magnons can be recovered, much in the same manner as for phonons. Interactions between magnons will broaden these peaks. Since the resulting line width is related to the decay rate of magnons, the magnon lifetime can be determined from scattering experiments.

15.5 Low-Dimensional Magnetic Systems

A lot of attention has been recently devoted to the study of magnetic systems in which the exchange interaction between atoms carrying magnetic moments

is appreciable only in one or two directions – while in other directions the energy of this type of interaction is negligibly small compared to the thermal energy. As far as magnetic properties are concerned, such systems should be considered as one- or two-dimensional, even if they behave as truly three-dimensional crystals in other respects.

15.5.1 Destruction of Magnetic Order by Thermal and Quantum Fluctuations

When studying the temperature dependence of magnetization in ferromagnets we found in (15.2.32) that the deviation from the saturation value is given by the integral

$$\Delta M = g\mu_{\mathrm{B}} \int \frac{\mathrm{d}\boldsymbol{k}}{(2\pi)^3} \frac{1}{\mathrm{e}^{\beta\hbar\omega_{\boldsymbol{k}}} - 1} \,. \tag{15.5.1}$$

Obviously, in the more general case of d dimensions, one has to evaluate the integral

$$\Delta M = g\mu_{\mathrm{B}} \int \frac{\mathrm{d}^d k}{(2\pi)^d} \frac{1}{\mathrm{e}^{\beta\hbar\omega_{\boldsymbol{k}}} - 1} \,. \tag{15.5.2}$$

To provide an estimate, suppose that the dispersion relation is strictly quadratic. At low temperatures, where only low-energy magnons are excited, this is a good approximation. Just as in the method employed for phonons, we shall integrate over a d-dimensional sphere of the same volume as the Brillouin zone. This leads to

$$\Delta M = g\mu_{\mathrm{B}} \int \frac{\mathrm{d}^d k}{(2\pi)^d} \frac{1}{\mathrm{e}^{\beta D k^2} - 1} = g\mu_{\mathrm{B}} K_d \int_0^{k_{\max}} \frac{k^{d-1}}{\mathrm{e}^{\beta D k^2} - 1} \, \mathrm{d}k \,, \tag{15.5.3}$$

where $K_d = 1/(\pi^{d/2} 2^{d-1} \Gamma(d/2))$. In terms of the new variable $x = \beta D k^2$ the integral becomes

$$\Delta M = \tfrac{1}{2} g\mu_{\mathrm{B}} K_d \left(\frac{k_{\mathrm{B}} T}{D}\right)^{d/2} \int_0^{x_{\max}} \frac{x^{(d-2)/2}}{\mathrm{e}^x - 1} \, \mathrm{d}x \,. \tag{15.5.4}$$

In $d > 2$ dimensions the integral is convergent. Its value can be determined approximately by choosing the upper limit as infinity, and using the formulas given in Appendix C. For $d \leq 2$, on the other hand, it is readily seen from the expansion of the integrand about $x = 0$ that the integral

$$\int_0^{\infty} \frac{x^{(d-2)/2}}{\mathrm{e}^x - 1} \, \mathrm{d}x \sim \int_0^{\infty} \frac{x^{(d-2)/2}}{x} \, \mathrm{d}x \tag{15.5.5}$$

blows up at the lower limit. This indicates that the spin-wave approximation cannot be applied to such systems. An even stronger statement can also be

made: long-range ferromagnetic order cannot exist at any finite temperature in the isotropic Heisenberg model for $d \leq 2$. The ordered ferromagnetic state can appear only as the ground state at $T = 0$, as at arbitrarily low but finite temperatures thermal fluctuations destroy the order, and spin–spin correlation functions decay exponentially.

For antiferromagnets a somewhat different calculation has to be performed. According to (15.3.23), the sublattice magnetization has to be determined from the expression

$$M_{\mathrm{A}}(T) = M_{\mathrm{A}}(0) - g\mu_{\mathrm{B}} \frac{1}{V} \sum_{\boldsymbol{k}} \left[u_{\boldsymbol{k}}^2 \langle \alpha_{\boldsymbol{k}}^\dagger \alpha_{\boldsymbol{k}} \rangle + v_{\boldsymbol{k}}^2 \langle \beta_{\boldsymbol{k}}^\dagger \beta_{\boldsymbol{k}} \rangle \right] . \tag{15.5.6}$$

Using the Bose–Einstein statistics for the magnon occupation number, and exploiting the fact that the thermal correction is governed mostly by long-wavelength magnons, the sum in the above expression can be approximated by the integral

$$\frac{1}{V} \sum_{\boldsymbol{k}} \frac{1}{\mathrm{e}^{\beta\hbar\omega_{\boldsymbol{k}}} - 1} \frac{1}{\sqrt{1 - \gamma_{\boldsymbol{k}}^2}} \sim \int \frac{\mathrm{d}^d k}{(2\pi)^d} \frac{1}{\mathrm{e}^{\beta D k} - 1} \frac{1}{k} . \tag{15.5.7}$$

For any finite temperature there is a region close to the lower limit of integration where $\beta D k < 1$. To determine the contribution of this region, one has to evaluate

$$K_d \int \frac{k^{d-1}\, \mathrm{d}k}{\beta D k^2} . \tag{15.5.8}$$

At the lower limit ($k = 0$) the last integral is finite for $d = 3$, while it is logarithmically divergent for $d = 2$. The divergence is even stronger for $d < 2$. This shows that antiferromagnetic order is also destroyed by thermal fluctuations in low-dimensional systems. This is in agreement with the Mermin–Wagner theorem, which has already been mentioned in connection with thermal disordering in two-dimensional lattices. The theorem, which was originally formulated for models described by the Heisenberg Hamiltonian, declares that in $d \leq 2$ dimensions no long-range ordered state may exist at any finite temperature that breaks a continuous symmetry of the Hamiltonian. This clearly applies to the isotropic Heisenberg model, where the continuous symmetry is the rotational symmetry of the spins.

The problem of the ground state is even more interesting. Of course, an ordered ferromagnetic ground state may exist at $T = 0$ in arbitrary dimension because it is an eigenstate of the Hamiltonian (in other words: because the order parameter of the ferromagnetic state, magnetization, is conserved). The situation is different in antiferromagnets, where the Néel state is not an eigenstate. Long-range order may exist at $T = 0$ in two-dimensional systems, but it is destroyed by quantum fluctuations even at $T = 0$ for $d < 2$.

To demonstrate this, consider the zero-point spin contraction that characterizes the correction to the Néel state. According to (15.3.21), the correction to the average value of the spin due to spin waves is

$$\Delta S^z = \frac{1}{N}\sum_{\boldsymbol{k}}\left[\frac{1}{\sqrt{1-\gamma_{\boldsymbol{k}}^2}}-1\right] \propto \frac{V}{N}\int\frac{\mathrm{d}^d k}{(2\pi)^d}\frac{1}{k}\,. \tag{15.5.9}$$

For a two-dimensional square lattice, the correction is about 0.2. Thus, a spin $S = 1/2$ is contracted to approximately 0.3, nonetheless the Néel-type order is preserved. For $d < 2$, and in particular for $d = 1$ the contribution at the lower limit is divergent, indicating that Néel-type antiferromagnetic order cannot exist in the ground state of one-dimensional systems. As we shall see, the ground state of the one-dimensional isotropic antiferromagnetic Heisenberg chain can be regarded as a singlet spin liquid.

15.5.2 Vortices in the Two-Dimensional Planar Model

The foregoing considerations were concerned with systems described by a Heisenberg model that is isotropic in spin space. As it was shown, such systems are on a borderline when the spatial dimension of the lattice of spins is two. In higher dimensions, e.g. in three-dimensional systems an ordered magnetic state – characterized by some order parameter – can emerge at low temperatures, and transition to the disordered phase takes place at a finite critical point. The correlation length becomes infinite in this point, and the spin–spin correlation function exhibits a power-law decay. At all other temperatures the correlation function decays exponentially.

On the other hand, we have seen in two dimensions that an ordered ground state may emerge for antiferromagnetic and ferromagnetic couplings alike, however, in accordance with the Mermin–Wagner theorem, there does not exist any state at finite temperature that breaks the continuous rotational symmetry of the Heisenberg model and features long-range order, since thermal fluctuations destroy any such order. The critical (Curie or Néel) temperature of the isotropic Heisenberg model is $T_\mathrm{c} = 0$ for $d = 2$. The spin–spin correlation function decays exponentially at any finite temperature.

The Mermin–Wagner theorem does not apply to the Ising model, for the latter does not possess continuous rotational symmetry in spin space. The transition between ordered and disordered phases is well known to occur at finite temperature in the two-dimensional Ising model. What about the planar, or XY model that falls between the Ising model and the isotropic Heisenberg model with respect to the dimensionality of the allowed spin space? In this model, by definition, spins lie in the (x, y) plane. As we shall see, this model exhibits an unusual phase transition. It takes place at a finite temperature, hence quantum effects may be ignored. Regarding the spins as classical vectors, it will be assumed that these vectors are of unit length, located at lattice sites $\boldsymbol{R}_i$, that they lie in the (x, y) plane, and their orientation is characterized by a polar angle θ_i:

$$S_i^x = \cos\theta_i\,, \qquad S_i^y = \sin\theta_i\,. \tag{15.5.10}$$

Assuming that only nearest-neighbor spins interact, the Hamiltonian of the XY model reads

$$\mathcal{H} = -J \sum_{<ij>} \left(\cos\theta_i \cos\theta_j + \sin\theta_i \sin\theta_j \right) = -J \sum_{<ij>} \cos(\theta_i - \theta_j), \quad (15.5.11)$$

where $\langle ij \rangle$ denotes adjacent lattice sites. Since the model has a continuous rotation symmetry in the (x, y) plane, at finite temperatures there cannot exist a phase with long-range order that breaks this symmetry. If, nevertheless, there exists an "ordered" phase at low temperatures, it cannot be truly ordered – however the correlation in the orientation of the classical vectors may be stronger than in a usual disordered state. The phase transition manifests itself in a different analytic form of the correlation functions at low and high temperatures.

At finite temperatures the free energy of the system is obtained from the partition function

$$Z = \int \prod_i \mathrm{d}\theta_i \, \mathrm{e}^{-\mathcal{H}/k_\mathrm{B}T}, \quad (15.5.12)$$

which is the sum (or integral) over all possible configurations of the angles θ_i. Some of the configurations correspond to spin-wave-like excitations. To study their effect, we shall assume that the spins of nearest neighbors are only slightly rotated with respect to one another, and thus it is sufficient to keep the leading-order term in the expansion of the cosine:

$$\mathcal{H} = E_0 + \tfrac{1}{2} J \sum_{<ij>} (\theta_i - \theta_j)^2 \,, \quad (15.5.13)$$

where E_0 is the energy of the completely ordered state. Replacing the variable θ_i defined at discrete lattice sites by the smooth function $\theta(\boldsymbol{r})$, we have

$$\mathcal{H} = E_0 + \tfrac{1}{2} J \int \mathrm{d}\boldsymbol{r} \left[\boldsymbol{\nabla} \theta(\boldsymbol{r}) \right]^2 . \quad (15.5.14)$$

In terms of the Fourier components:

$$\mathcal{H} = E_0 + \tfrac{1}{2} J \sum_{\boldsymbol{k}} k^2 |\hat{\theta}(\boldsymbol{k})|^2 \,. \quad (15.5.15)$$

However, we may encounter other configurations in which – apart from some singular points – θ_i varies slightly between adjacent lattice sites, nevertheless the sum of the differences $\Delta\theta_i$ along a closed path C encircling a singular point will not be zero but an integral multiple of 2π,

$$\sum_C \Delta\theta_i = 2\pi q \,. \quad (15.5.16)$$

Such a configuration, featuring a singular point, is obtained, e.g., for

$$\theta_i = \theta_0 + q\varphi_i = \theta_0 + q \arctan \frac{y_i}{x_i} \,, \quad (15.5.17)$$

that is when the inclination of a spin is q times the polar angle of its lattice site $\boldsymbol{r}_i$. Needless to say, θ_i may just as well be a more complicated function of the polar angle. Such configurations, with quantum number $q = \pm 1$ and $q = \pm 2$ are presented in Figs. 15.4 and 15.5, respectively. Because of their vortex-like character, such configurations are called *vortices* and the quantum number is called vorticity.

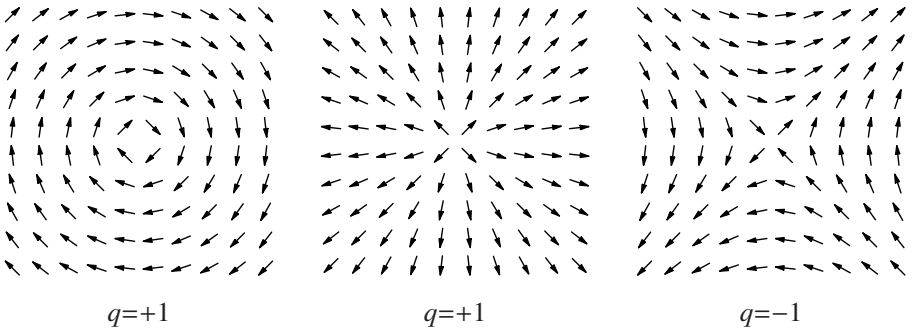


Fig. 15.4. Vortex configurations of quantum number $q = \pm 1$ in the two-dimensional planar model

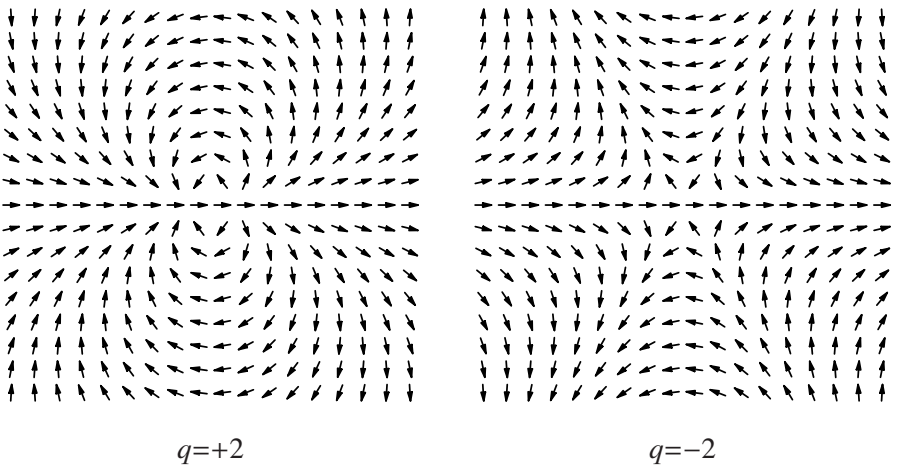


Fig. 15.5. Vortex configurations of quantum number $q = \pm 2$ in the two-dimensional planar model

Following the procedure used to show the topological character of dislocations it is easy to show that the integer q is a topological quantum number.

In this classical spin model the variable $\theta(\boldsymbol{r})$ varies on the unit circle, thus the order parameter space is the unit circle. Going counter-clockwise around the core of a vortex of vorticity $q = 1$, the variable $\theta(\boldsymbol{r})$ goes around the unit circle once. If the spin orientation is distorted continuously, i.e., $\theta(\boldsymbol{r})$ is varied continuously locally, then every time when one goes around the vortex along the same circuit, a deformed path is found in the order parameter space, but its deformation is also continuous, and since it has to remain on the unit circle, it is not possible to reduce it to zero or to reach configurations of another quantum number.

Even when studying the role of vortex configurations it is admissible – apart from the immediate neighborhood of the core of the vortex – to approximate θ_i defined at lattice points only by a continuous function $\theta(\boldsymbol{r})$. Then, going around the vortex of vorticity q on a closed path C, we have, instead of (15.5.16),

$$\oint_C \frac{\mathrm{d}\theta(\boldsymbol{r})}{\mathrm{d}\boldsymbol{r}} \mathrm{d}\boldsymbol{l} = 2\pi q \,. \tag{15.5.18}$$

Assuming that the rate of change of the angle is constant on a circle of radius r around the core,

$$|\boldsymbol{\nabla}\theta(\boldsymbol{r})| = \frac{2\pi q}{2\pi r} \,. \tag{15.5.19}$$

In polar coordinates

$$\boldsymbol{\nabla}\theta(\boldsymbol{r}) = \frac{q}{r}\boldsymbol{e}_\varphi \,, \tag{15.5.20}$$

where $\boldsymbol{e}_\varphi$ is the tangential unit vector drawn to the point $\boldsymbol{r}$. Using the radial unit vector $\boldsymbol{e}_r$ and the unit vector $\boldsymbol{e}_z$ perpendicular to the plane,

$$\boldsymbol{\nabla}\theta(\boldsymbol{r}) = -\frac{q}{r}\boldsymbol{e}_r \times \boldsymbol{e}_z = -q\boldsymbol{\nabla} \times (\boldsymbol{e}_z \ln r) \,. \tag{15.5.21}$$

To estimate the energy of such a configuration, we shall further assume that this form is valid all the way from the lattice constant a to a radius R characteristic of the size of the system. Then

$$\begin{aligned} E &= \tfrac{1}{2}J \int \left[\boldsymbol{\nabla}\theta(\boldsymbol{r})\right]^2 \mathrm{d}\boldsymbol{r} = \tfrac{1}{2}J \int_a^R \left[\boldsymbol{\nabla}\theta(\boldsymbol{r})\right]^2 2\pi r \,\mathrm{d}r \\ &= \pi J \int_a^R \frac{q^2}{r} \mathrm{d}r = \pi J q^2 \ln \frac{R}{a} \,. \end{aligned} \tag{15.5.22}$$

Since this energy diverges as the size of the system is increased indefinitely, no such configuration is excited thermally at low temperatures.

However, such configurations may appear at high temperatures, since they may lower the free energy through the $-TS$ term. As the vortex core can be at any lattice site, the number of vortex configurations is R^2/a^2, and the entropy of the vortex is

$$S = k_{\mathrm{B}} \ln \left(\frac{R}{a} \right)^2 . \tag{15.5.23}$$

Consequently, the free energy of a system with a single vortex is given by

$$F = E - TS = (\pi J q^2 - 2k_{\mathrm{B}}T) \ln(R/a) . \tag{15.5.24}$$

This indicates that individual vortices of vorticity q may appear spontaneously above $T = \frac{1}{2}\pi J q^2/k_{\mathrm{B}}$. As it was pointed out by Berezinskii[8] and later by Kosterlitz and Thouless,[9] the behavior of the system – i.e., the character of the spin–spin correlations – changes at the temperature where the first free vortices appear,

$$T_{\mathrm{BKT}} = \tfrac{1}{2}\pi J/k_{\mathrm{B}} . \tag{15.5.25}$$

This is the *BKT (Berezinskii–Kosterlitz–Thouless) transition.*

At high temperatures – as it is usual in disordered phases –, the correlation function decays exponentially with distance due to the disordered spatial distribution of vortices. The correlation length $\xi(T)$ characterizing the decay increases as the temperature is lowered, and diverges at T_{BKT}. At this temperature the decay of the correlation function is no longer exponential but power-law-like, as usual in a critical point. The particular feature of the BKT transition is that this power-law behavior survives even below T_{BKT}, in the low-temperature regime, where thermal fluctuations prevent ordering, but with a temperature-dependent exponent:

$$\Gamma(r) \propto \left(\frac{1}{r} \right)^{\eta(T)} . \tag{15.5.26}$$

In the simplest approximation, where only the effects of spin-wave-like harmonic fluctuations are taken into account, $\eta(T) = k_{\mathrm{B}}T/(2\pi J)$. To demonstrate this, consider the correlation function

$$\begin{aligned} \langle \boldsymbol{S}(0) \cdot \boldsymbol{S}(\boldsymbol{r}) \rangle &= \langle \cos(\theta(0) - \theta(\boldsymbol{r})) \rangle \\ &= \tfrac{1}{2} \Big[\big\langle \mathrm{e}^{\mathrm{i}(\theta(0)-\theta(\boldsymbol{r}))} \big\rangle + \big\langle \mathrm{e}^{-\mathrm{i}(\theta(0)-\theta(\boldsymbol{r}))} \big\rangle \Big] . \end{aligned} \tag{15.5.27}$$

If spin waves alone are taken into account – that is, if we employ the harmonic approximation and calculate the thermal average with the weight factor from the Hamiltonian (15.5.15) – then, according to (13.3.11), the averaging procedure can be performed in the exponent,

$$\langle \boldsymbol{S}(0) \cdot \boldsymbol{S}(\boldsymbol{r}) \rangle = \exp \Big[-\tfrac{1}{2} \big\langle (\theta(0) - \theta(\boldsymbol{r}))^2 \big\rangle \Big] . \tag{15.5.28}$$

Writing the exponent in terms of Fourier components,

[8] V. L. Berezinskii, 1970.
[9] J. M. Kosterlitz and D. J. Thouless, 1972.

$$\tfrac{1}{2}\big\langle(\theta(0)-\theta(\boldsymbol{r}))^2\big\rangle = \sum_{\boldsymbol{k}}(1-\cos\boldsymbol{k}\cdot\boldsymbol{r})\big\langle|\theta(\boldsymbol{k})|^2\big\rangle\,, \tag{15.5.29}$$

and using the Boltzmann weight factor that follows from (15.5.15), we have

$$\tfrac{1}{2}\big\langle(\theta(0)-\theta(\boldsymbol{r}))^2\big\rangle = \sum_{\boldsymbol{k}}(1-\cos\boldsymbol{k}\cdot\boldsymbol{r})\,\frac{k_{\mathrm{B}}T}{Jk^2}\,. \tag{15.5.30}$$

Replacing the sum over the Brillouin zone by an integral over a circle of radius π/a, for large values of r

$$\tfrac{1}{2}\big\langle(\theta(0)-\theta(\boldsymbol{r}))^2\big\rangle = \frac{k_{\mathrm{B}}T}{2\pi J}\ln(\pi r/a)\,. \tag{15.5.31}$$

Substituting this into the exponent, as required by (15.5.28), the form (15.5.26) is indeed recovered for the correlation function. Below the transition point, throughout the temperature range from $T=0$ up to T_{BKT} the system behaves as it were critical, however, the critical exponents are not universal but depend on the coupling and the temperature. This is the so-called *Berezinskii–Kosterlitz–Thouless phase* (BKT phase).

Although free vortices appear only above the transition point, configurations that can be regarded as bound states of two oppositely "charged" vortices may exist at lower temperatures as well. Such configurations are shown in Fig. 15.6.

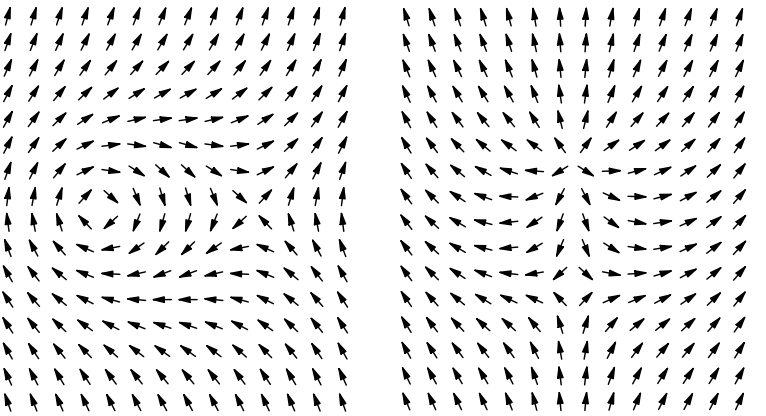


Fig. 15.6. Configurations corresponding to the bound state of vortices of opposite vorticity

Assume that the system contains only two vortices, at $\boldsymbol{r}_1$ and at $\boldsymbol{r}_2$, and that their quantum numbers are q_1 and q_2. The angular variable $\theta(\boldsymbol{r})$ can be chosen in the form

$$\theta(\boldsymbol{r}) = q_1 \arctan \frac{y - y_1}{x - x_1} + q_2 \arctan \frac{y - y_2}{x - x_2} \,. \tag{15.5.32}$$

The path integral of $\boldsymbol{\nabla}\theta(\boldsymbol{r})$ around a closed circuit gives $2\pi q_1$, $2\pi q_2$ or $2\pi(q_1 + q_2)$, depending on whether the circuit goes around the first, the second or both vortices.

As will be shown later (see (15.5.54)), the energy of this configuration is

$$\tfrac{1}{2}J \int \mathrm{d}\boldsymbol{r} \big[\nabla\theta(\boldsymbol{r})\big]^2 = E_1 + E_2 + 2\pi J q_1 q_2 \ln \frac{R}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|} \,, \tag{15.5.33}$$

where E_1 and E_2 are the energies of the individual vortices, as given in (15.5.22). The third term, the interaction energy diverges logarithmically as the size of the system increases. Writing the total energy in the form

$$\tfrac{1}{2}J \int \mathrm{d}\boldsymbol{r} \big[\nabla\theta(\boldsymbol{r})\big]^2 = \pi J (q_1 + q_2)^2 \ln \frac{R}{a} - 2\pi J q_1 q_2 \ln \frac{|\boldsymbol{r}_1 - \boldsymbol{r}_2|}{a} \,, \tag{15.5.34}$$

the size-dependent divergent term is seen to disappear when the two vortices are of opposite vorticity, $q_1 + q_2 = 0$. The energy depends logarithmically on the distance between the two vortex cores, and it might be much smaller than the energy of two individual vortices, especially if the distance of the cores is on the order of the lattice constant. If this energy is small enough, a pair of vortices with opposite vorticities may be thermally excited. Such vortex pairs are indeed observed in the low-temperature Berezinskii–Kosterlitz–Thouless phase of the planar XY model, where the system is "neutral" in the sense that the total vorticity is zero.

Thus, below T_{BKT}, spin-wave-like configurations and vortex pairs are present simultaneously. To determine the total energy we assume that the density of vortices is given by

$$\rho(\boldsymbol{r}) = \sum_i q_i \delta(\boldsymbol{r} - \boldsymbol{r}_i) \,. \tag{15.5.35}$$

The generalization of (15.5.18) for this case is

$$\oint_C \boldsymbol{\nabla}\theta(\boldsymbol{r}) \cdot \mathrm{d}\boldsymbol{l} = 2\pi \sum_i q_i \,, \tag{15.5.36}$$

where the sum is over all vortices inside the closed path C. Transforming, by means of Stokes' theorem, the line integral into an integral over the region enclosed by C,

$$\int \boldsymbol{e}_z \cdot \big(\boldsymbol{\nabla} \times \boldsymbol{\nabla}\theta(\boldsymbol{r})\big) \mathrm{d}\boldsymbol{r} = 2\pi \sum_i q_i \,. \tag{15.5.37}$$

This can be satisfied for any closed region if

$$\boldsymbol{\nabla} \times \boldsymbol{\nabla}\theta(\boldsymbol{r}) = 2\pi \boldsymbol{e}_z \sum_i q_i \delta(\boldsymbol{r} - \boldsymbol{r}_i) = 2\pi \boldsymbol{e}_z \rho(\boldsymbol{r}) \,. \tag{15.5.38}$$

Seeking the solution in the form

$$\theta(\boldsymbol{r}) = \theta_0(\boldsymbol{r}) + \psi(\boldsymbol{r})\,, \tag{15.5.39}$$

where $\psi(\boldsymbol{r})$ is assumed to be a continuous, single-valued function, its gradient is vortex-free, i.e., $\boldsymbol{\nabla} \times \boldsymbol{\nabla}\psi(\boldsymbol{r}) = 0$, and

$$\oint_C \boldsymbol{\nabla}\psi(\boldsymbol{r}) \cdot \mathrm{d}\boldsymbol{l} = 0\,. \tag{15.5.40}$$

Thus $\psi(\boldsymbol{r})$ describes the fluctuations in the orientation of the spin vectors due to spin waves, while $\theta_0(\boldsymbol{r})$ comes from the vortices or vortex pairs. If $\theta_0(\boldsymbol{r})$ is chosen so that

$$\boldsymbol{\nabla}\theta_0(\boldsymbol{r}) = -\boldsymbol{\nabla} \times \big(\boldsymbol{e}_z \chi(\boldsymbol{r})\big) \tag{15.5.41}$$

is satisfied, then

$$\boldsymbol{\nabla} \times \boldsymbol{\nabla}\theta_0(\boldsymbol{r}) = \boldsymbol{e}_z \boldsymbol{\nabla}^2 \chi(\boldsymbol{r})\,, \tag{15.5.42}$$

which implies that $\chi(\boldsymbol{r})$ satisfies the two-dimensional Poisson equation

$$\boldsymbol{\nabla}^2 \chi(\boldsymbol{r}) = 2\pi \sum_i q_i \delta(\boldsymbol{r} - \boldsymbol{r}_i) = 2\pi\rho(\boldsymbol{r})\,. \tag{15.5.43}$$

The solution of this equation can be written as the sum of two-dimensional Coulomb potentials:

$$\chi(\boldsymbol{r}) = \sum_i q_i \ln(|\boldsymbol{r} - \boldsymbol{r}_i|)\,. \tag{15.5.44}$$

A more precise form is obtained in terms of the Green function of the Laplacian, which satisfies the equation

$$\boldsymbol{\nabla}^2 g(\boldsymbol{r}) = \delta(\boldsymbol{r})\,, \tag{15.5.45}$$

which leads to

$$\chi(\boldsymbol{r}) = 2\pi \int \mathrm{d}\boldsymbol{r}'\, g(\boldsymbol{r} - \boldsymbol{r}')\rho(\boldsymbol{r}')\,. \tag{15.5.46}$$

In two dimensions the asymptotic solution for the Green function for large distances from the core of the vortex is:

$$g(\boldsymbol{r}) = -\int \frac{\mathrm{d}\boldsymbol{k}}{(2\pi)^2} \frac{\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}}{k^2} \approx \frac{1}{2\pi} \ln\left(|\boldsymbol{r}|/a\right) - \frac{1}{2\pi} \ln\left(R/a\right)\,, \tag{15.5.47}$$

where R is the radius of the sample.

The energy of the configuration that contains vortices as well can be determined using (15.5.14),

$$\begin{aligned} E &= E_0 + \tfrac{1}{2}J \int \mathrm{d}\boldsymbol{r} \big[\boldsymbol{\nabla}\psi(\boldsymbol{r}) + \boldsymbol{\nabla}\theta_0(\boldsymbol{r})\big]^2 \\ &= E_0 + \tfrac{1}{2}J \int \mathrm{d}\boldsymbol{r} \big[\boldsymbol{\nabla}\psi(\boldsymbol{r})\big]^2 + \tfrac{1}{2}J \int \mathrm{d}\boldsymbol{r} \big[\boldsymbol{\nabla}\theta_0(\boldsymbol{r})\big]^2\,, \end{aligned} \tag{15.5.48}$$

since the mixed term containing the integral of $\boldsymbol{\nabla}\psi(\boldsymbol{r}) \cdot \big(\boldsymbol{\nabla} \times (\boldsymbol{e}_z \chi(\boldsymbol{r}))\big)$ vanishes, as the reader may verify by a simple integration by parts. This means that the energy contributions of spin waves and vortices are independent of each other. Below we shall only deal with the contribution of vortices. Since (15.5.41),

$$\boldsymbol{\nabla}\theta_0(\boldsymbol{r}) = \left(-\frac{\partial \chi(\boldsymbol{r})}{\partial y} \, \frac{\partial \chi(\boldsymbol{r})}{\partial x} \right) , \tag{15.5.49}$$

implies

$$\big(\boldsymbol{\nabla}\theta_0(\boldsymbol{r})\big)^2 = \big(\boldsymbol{\nabla}\chi(\boldsymbol{r})\big)^2 , \tag{15.5.50}$$

the energy of vortices can be obtained from

$$E = E_0 + \tfrac{1}{2} J \int \mathrm{d}\boldsymbol{r} \big[\boldsymbol{\nabla}\chi(\boldsymbol{r})\big]^2 , \tag{15.5.51}$$

or alternatively, after an integration by parts, from

$$E = E_0 - \tfrac{1}{2} J \int \mathrm{d}\boldsymbol{r} \chi(\boldsymbol{r}) \boldsymbol{\nabla}^2 \chi(\boldsymbol{r}) . \tag{15.5.52}$$

Using the expressions obtained for $\chi(\boldsymbol{r})$ and $\boldsymbol{\nabla}^2\chi(\boldsymbol{r})$,

$$E = E_0 - 2\pi^2 J \iint \mathrm{d}\boldsymbol{r} \, \mathrm{d}\boldsymbol{r}' \rho(\boldsymbol{r}) g(\boldsymbol{r} - \boldsymbol{r}') \rho(\boldsymbol{r}') . \tag{15.5.53}$$

From the explicit form of the Green function and assumption (15.5.35) for the density of vortices

$$E = E_0 - \pi J \sum_{i \neq j} q_i q_j \ln \left| \frac{\boldsymbol{r}_i - \boldsymbol{r}_j}{a} \right| + \pi J \left(\sum_i q_i \right)^2 \ln \frac{R}{a} , \tag{15.5.54}$$

in agreement the expression anticipated for the energy of a vortex pair. The term depending on the size of the sample vanishes if $\sum_i q_i = 0$, i.e., if the total vorticity is zero.

In the foregoing discussion the correlation function in the low-temperature phase and the transition temperature were determined under the assumption that only spin waves are excited below T_{BKT}. A more precise treatment requires the inclusion of vortex pairs as well. A rather tedious calculation, whose details cannot be given here, shows that the effects of the vortex pairs can be absorbed into the parameters of the vortex-free model. The behavior of the full system is then similar to a simpler one in which only spin waves are present but the parameters are modified, renormalized. Below the transition point the correlation function shows a power-law decay but the exponent $\eta(T)$ is renormalized. The transition temperature itself is also renormalized. The first corrections coming from the vortex pairs give

$$\frac{\pi J}{2 k_{\mathrm{B}} T_{\mathrm{BKT}}} \approx 1 + 1.3\pi \mathrm{e}^{-\pi^2 J / 2 k_{\mathrm{B}} T_{\mathrm{BKT}}} . \tag{15.5.55}$$

The most accurate Monte-Carlo simulations give a much lower critical temperature, $k_{\mathrm{B}}T_{\mathrm{BKT}} = 0.893J$. Nevertheless, the critical exponent η at the true transition point is $\eta(T_{\mathrm{BKT}}) = 1/4$ – the same as the result obtained when vortex pairs were neglected.

While for increasing temperatures more and more vortex pairs are excited thermally in the Berezinskii–Kosterlitz–Thouless phase, they are all confined to neutral pairs made up of nearby vortices. Above the transition point they become deconfined (liberated). The neutral bound pairs disintegrate into freely moving individual vortices. The correlation function changes character: from this point it decays exponentially. It can also be shown that the finite correlation length ξ is not proportional to some inverse power of $t = (T - T_{\mathrm{BKT}})/T_{\mathrm{BKT}}$, as it occurs above ordinary critical points but it depends exponentially on t:

$$\xi \propto \exp(bt^{-1/2})\,. \tag{15.5.56}$$

15.5.3 The Spin-1/2 Anisotropic Ferromagnetic Heisenberg Chain

It was mentioned earlier that the energy of states with two raised or lowered spins relative to the ground state could, in principle, be calculated in systems of arbitrary dimensionality and spin. The exact determination of further states, involving more flipped spins, is impossible for general S, even in one-dimensional spin chains. The spin-1/2 Heisenberg chain is particular, for an exact determination of excited states with more reversed spins is possible, and the calculation can be easily extended to the anisotropic case, too.

Using the Pauli operators instead of the spin operators, and separating the energy of the ferromagnetic state, the Hamiltonian is customarily cast in the form

$$\mathcal{H} = -J\sum_l \left[\tfrac{1}{2}\left(\sigma_l^+\sigma_{l+1}^- + \sigma_l^-\sigma_{l+1}^+\right) + \tfrac{1}{4}\Delta\left(\sigma_l^z\sigma_{l+1}^z - 1\right)\right] - \tfrac{1}{4}NJ\Delta\,, \tag{15.5.57}$$

where again, similarly to (15.2.77) a factor of two has been dropped. When $J > 0$, $\Delta = 1$ corresponds to an isotropic ferromagnet, and $\Delta = 0$ to the pure planar model. $J < 0$ would lead to antiferromagnets, but via the rotation of every second spin through 180° about the z-axis, this case is found to be equivalent to $J > 0$, $\Delta < 0$. Consequently, the parameters $J > 0$, $\Delta = -1$ and $J < 0$, $\Delta = 1$ can equally be used in the discussion of isotropic antiferromagnets.

Taking again the state with all spins aligned upward as the ferromagnetic ground state, we shall seek eigenstates with one reversed spin in the form

$$|\Psi\rangle = \sum_l c(x_l)\sigma_l^-|F\rangle\,. \tag{15.5.58}$$

When the Hamiltonian acts on this state, the following eigenvalue equation is obtained:

$$Ec(x_l) = -\tfrac{1}{4}NJ\Delta c(x_l) + \tfrac{1}{2}J\left[2\Delta c(x_l) - c(x_{l+1}) - c(x_{l-1})\right] . \qquad (15.5.59)$$

We shall try to find the solutions using the trial function $c(x_l) = A\mathrm{e}^{\mathrm{i}kx_l}$. The periodic boundary condition implies $c(x_l) = c(x_l + Na)$, hence

$$k = \frac{2\pi}{Na}I, \qquad \text{where} \qquad I = 0, \pm 1, \pm 2, \ldots, N/2 . \qquad (15.5.60)$$

Substitution of this formula into the eigenvalue equation leads to

$$E = -\tfrac{1}{4}NJ\Delta + J\left(\Delta - \cos ka\right) , \qquad (15.5.61)$$

thus the excitation energy is

$$\hbar\omega_k = J\left(\Delta - \cos ka\right) . \qquad (15.5.62)$$

In the isotropic case, the spin-wave spectrum starting off as k^2 is recovered. For $\Delta > 1$, the excitation spectrum features a finite gap. For $\Delta < 1$, on the other hand, the energy of the long-wavelength excitations turns out to be negative, indicating that the ferromagnetic state is no longer the ground state. We shall determine the true ground state later.

When two spins are reversed, the wavefunction is chosen as

$$|\Psi_2\rangle = \sum_{ll'} c(x_l, x_{l'})\sigma_l^- \sigma_{l'}^- |F\rangle , \qquad (15.5.63)$$

in line with (15.2.82), and $x_l < x_{l'}$ is assumed. This restriction is necessary as in a spin-1/2 chain the spin cannot be reversed twice at the same lattice site. When the two reversed spins are not at adjacent lattice sites, the Schrödinger equation yields the following relation for the coefficients:

$$\begin{aligned} E\,c(x_l, x_{l'}) = &-\tfrac{1}{4}NJ\Delta c(x_l, x_{l'}) + 2J\Delta c(x_l, x_{l'}) \\ &- \tfrac{1}{2}Jc(x_{l-1}, x_{l'}) - \tfrac{1}{2}Jc(x_{l+1}, x_{l'}) \\ &- \tfrac{1}{2}Jc(x_l, x_{l'-1}) - \tfrac{1}{2}Jc(x_l, x_{l'+1}) . \end{aligned} \qquad (15.5.64)$$

Writing the solution in the plane-wave-like form

$$c(x_l, x_{l'}) = A_{12}\mathrm{e}^{\mathrm{i}(k_1 x_l + k_2 x_{l'})} + A_{21}\mathrm{e}^{\mathrm{i}(k_2 x_l + k_1 x_{l'})} , \qquad (15.5.65)$$

the equation for the energy eigenvalue becomes

$$E = -\tfrac{1}{4}NJ\Delta + J\left(\Delta - \cos k_1 a\right) + J\left(\Delta - \cos k_2 a\right) . \qquad (15.5.66)$$

The excitation energy is again equal to the sum of the energies of two free magnons, therefore these excitations are expected to form a continuum. To determine the possible values of k_1 and k_2, we have to examine the case when spins are reversed at two adjacent lattice sites. The following equation holds for the amplitude of such states:

$$\begin{aligned} E\,c(x_l, x_{l+1}) = &-\tfrac{1}{4}NJ\Delta c(x_l, x_{l+1}) + J\Delta c(x_l, x_{l+1}) \\ &- \tfrac{1}{2}Jc(x_{l-1}, x_{l+1}) - \tfrac{1}{2}Jc(x_l, x_{l+2})\,. \end{aligned} \tag{15.5.67}$$

By demanding that this equation also hold with the previously given forms of the wavefunction and the energy – (15.5.65) and (15.5.66) –, restrictions are imposed on the amplitudes A_{12} and A_{21}. To determine these, we shall assume that (15.5.64) formally holds in the $l' = l + 1$ case, too, that is,

$$\begin{aligned} E\,c(x_l, x_{l+1}) = &-\tfrac{1}{4}NJ\Delta c(x_l, x_{l+1}) + 2J\Delta c(x_l, x_{l+1}) \\ &- \tfrac{1}{2}Jc(x_{l-1}, x_{l+1}) - \tfrac{1}{2}Jc(x_{l+1}, x_{l+1}) \\ &- \tfrac{1}{2}Jc(x_l, x_l) - \tfrac{1}{2}Jc(x_l, x_{l+2})\,. \end{aligned} \tag{15.5.68}$$

When combined with the previous equation, consistency requires

$$2\Delta c(x_l, x_{l+1}) = c(x_l, x_l) + c(x_{l+1}, x_{l+1})\,. \tag{15.5.69}$$

Using the assumed form (15.5.65) for the coefficients,

$$2\Delta\left(A_{12}\mathrm{e}^{\mathrm{i}k_2 a} + A_{21}\mathrm{e}^{\mathrm{i}k_1 a}\right) = (A_{12} + A_{21})\left(1 + \mathrm{e}^{\mathrm{i}(k_1+k_2)a}\right)\,, \tag{15.5.70}$$

whence

$$\frac{A_{12}}{A_{21}} = -\frac{2\Delta\mathrm{e}^{\mathrm{i}k_1 a} - 1 - \mathrm{e}^{\mathrm{i}(k_1+k_2)a}}{2\Delta\mathrm{e}^{\mathrm{i}k_2 a} - 1 - \mathrm{e}^{\mathrm{i}(k_1+k_2)a}}\,. \tag{15.5.71}$$

Expressing the amplitudes in terms of a phase difference $\phi(k_1, k_2)$ defined on the interval $(-\pi, \pi)$,

$$A_{12} = A\,\mathrm{e}^{\mathrm{i}\phi(k_1,k_2)/2}\,, \qquad A_{21} = A\mathrm{e}^{-\mathrm{i}\phi(k_1,k_2)/2}\,, \tag{15.5.72}$$

which corresponds to choosing the coefficient $c(x_l, x_{l'})$ as

$$c(x_l, x_{l'}) = A\left(\mathrm{e}^{\mathrm{i}[k_1 x_l + k_2 x_{l'} + \phi(k_1,k_2)/2]} + \mathrm{e}^{\mathrm{i}[k_2 x_l + k_1 x_{l'} - \phi(k_1,k_2)/2]}\right)\,. \tag{15.5.73}$$

The two terms can be interpreted as follows. Two magnons propagate in the system. In their interaction – scattering by one another – the total wave number is not the only conserved quantity: k_1 and k_2 are conserved separately. On the other hand, when the two magnons pass through each other, the wavefunction undergoes a phase shift of $\phi(k_1, k_2)$.

From the expression

$$\mathrm{e}^{\mathrm{i}\phi(k_1,k_2)} = -\frac{2\Delta\mathrm{e}^{\mathrm{i}k_1 a} - 1 - \mathrm{e}^{\mathrm{i}(k_1+k_2)a}}{2\Delta\mathrm{e}^{\mathrm{i}k_2 a} - 1 - \mathrm{e}^{\mathrm{i}(k_1+k_2)a}}\,, \tag{15.5.74}$$

which is implied by (15.5.71) and (15.5.72),

$$\begin{aligned} \cot\tfrac{1}{2}\phi(k_1, k_2) &= \mathrm{i}\frac{\mathrm{e}^{\mathrm{i}\phi(k_1,k_2)} + 1}{\mathrm{e}^{\mathrm{i}\phi(k_1,k_2)} - 1} \\ &= \Delta\frac{\sin[(k_1 - k_2)a/2]}{\cos[(k_1 + k_2)a/2] - \Delta\cos[(k_1 - k_2)a/2]} \\ &= \Delta\frac{\cot(k_1 a/2) - \cot(k_2 a/2)}{(1 + \Delta) - (1 - \Delta)\cot(k_1 a/2)\cot(k_2 a/2)}\,. \end{aligned} \tag{15.5.75}$$

For the isotropic spin-1/2 Heisenberg model this is equivalent to the result (15.2.95).

The periodic boundary condition provides another relation between the as yet undetermined wave numbers and the phase. Just like in the calculation for the isotropic Heisenberg chain with arbitrary S, we arrive at equations (15.2.87) and (15.2.88) anew,

$$\mathrm{e}^{\mathrm{i}k_1 Na} = \mathrm{e}^{\mathrm{i}\phi(k_1,k_2)}, \qquad \mathrm{e}^{\mathrm{i}k_2 Na} = \mathrm{e}^{-\mathrm{i}\phi(k_1,k_2)}, \tag{15.5.76}$$

or equivalently

$$k_1 Na = 2\pi I_1 + \phi(k_1,k_2)\,, \qquad k_2 Na = 2\pi I_2 - \phi(k_1,k_2)\,, \tag{15.5.77}$$

where both I_1 and I_2 are integers such that $0 \le I_i < N$.

To get the excitation spectrum, the closed set of equations can be solved numerically or analytically in the large-N limit. It is found that the energies associated with real values of k_i fill almost continuously the region between the curves

$$\hbar\omega_-(k) = 2J(\Delta - \cos\tfrac{1}{2}ka) \quad \text{and} \quad \hbar\omega_+(k) = 2J(\Delta + \cos\tfrac{1}{2}ka)\,. \tag{15.5.78}$$

Apart from these solutions, others, associated with complex conjugate pairs ($k_2 = k_1^*$) arise, too, with an energy below the continuum,

$$\hbar\omega(k) = \frac{J}{\Delta}\left(\Delta - \cos\tfrac{1}{2}ka\right)\left(\Delta + \cos\tfrac{1}{2}ka\right), \tag{15.5.79}$$

where $k = k_1 + k_2 = 2\,\mathrm{Re}\,k_1$.

In the isotropic case the results obtained from (15.2.93) and (15.2.96) by setting $S = 1/2$ are recovered. It will prove convenient to introduce the variables

$$\lambda_1 = \cot(k_1 a/2) \qquad \text{and} \qquad \lambda_2 = \cot(k_2 a/2)\,, \tag{15.5.80}$$

called *rapidities*. Then making use of the ensuing formulas

$$\mathrm{e}^{\mathrm{i}k_1 a} = \frac{\lambda_1 + \mathrm{i}}{\lambda_1 - \mathrm{i}} \qquad \text{and} \qquad \mathrm{e}^{\mathrm{i}k_2 a} = \frac{\lambda_2 + \mathrm{i}}{\lambda_2 - \mathrm{i}}\,, \tag{15.5.81}$$

along with

$$\mathrm{e}^{\mathrm{i}\phi(k_1,k_2)} = \frac{\cot(\phi(k_1,k_2)/2) + \mathrm{i}}{\cot(\phi(k_1,k_2)/2) - \mathrm{i}}\,, \tag{15.5.82}$$

the boundary condition leads to the following algebraic equations for the rapidities

$$\left(\frac{\lambda_1 + \mathrm{i}}{\lambda_1 - \mathrm{i}}\right)^N = \frac{\lambda_1 - \lambda_2 + 2\mathrm{i}}{\lambda_1 - \lambda_2 - 2\mathrm{i}}\,, \qquad \left(\frac{\lambda_2 + \mathrm{i}}{\lambda_2 - \mathrm{i}}\right)^N = \frac{\lambda_2 - \lambda_1 + 2\mathrm{i}}{\lambda_2 - \lambda_1 - 2\mathrm{i}}\,. \tag{15.5.83}$$

It was first pointed out by BETHE[10] that in the special case of a spin-1/2 anisotropic Heisenberg chain the method above can be generalized for states with arbitrarily many reversed spins.

States with three reversed spins may be written as

$$|\Psi_3\rangle = \sum_{x_{l_1}<x_{l_2}<x_{l_3}} c(x_{l_1}, x_{l_2}, x_{l_3}) \big|\psi(x_{l_1}, x_{l_2}, x_{l_3})\big\rangle \,, \tag{15.5.84}$$

where

$$\big|\psi(x_{l_1}, x_{l_2}, x_{l_3})\big\rangle = \sigma_{l_1}^- \sigma_{l_2}^- \sigma_{l_3}^- |F\rangle \,. \tag{15.5.85}$$

Substituting this into the Schrödinger equation, we may follow the method used for states with two reversed spins – that is, separate equations have to be written for the cases when there are no neighbors and when there are neighbors among x_{l_1}, x_{l_2} and x_{l_3}. In analogy to (15.5.65), we shall use the ansatz

$$\begin{aligned} c(x_{l_1}, x_{l_2}, x_{l_3}) = {} & A_{123} \mathrm{e}^{\mathrm{i}(k_1 x_{l_1} + k_2 x_{l_2} + k_3 x_{l_3})} + A_{132} \mathrm{e}^{\mathrm{i}(k_1 x_{l_1} + k_3 x_{l_2} + k_2 x_{l_3})} \\ & + A_{213} \mathrm{e}^{\mathrm{i}(k_2 x_{l_1} + k_1 x_{l_2} + k_3 x_{l_3})} + A_{231} \mathrm{e}^{\mathrm{i}(k_2 x_{l_1} + k_3 x_{l_2} + k_1 x_{l_3})} \\ & + A_{312} \mathrm{e}^{\mathrm{i}(k_3 x_{l_1} + k_1 x_{l_2} + k_2 x_{l_3})} + A_{321} \mathrm{e}^{\mathrm{i}(k_3 x_{l_1} + k_2 x_{l_2} + k_1 x_{l_3})} \,, \end{aligned} \tag{15.5.86}$$

which indicates immediately that the total wave number of the state is $k = k_1 + k_2 + k_3$.

From the equations for nonadjacent lattice sites, the excitation energy

$$\hbar\omega = J\,(\Delta - \cos k_1 a) + J\,(\Delta - \cos k_2 a) + J\,(\Delta - \cos k_3 a) \tag{15.5.87}$$

emerges, regardless of the amplitude values – as if the energies of three independent magnons were summed. The allowed values of the wave numbers and the coefficients in the wavefunctions are determined by the boundary condition and the requirement that the same energy values have to satisfy the other types of equations, in which two or all three spin reversals take place at adjacent lattice sites. It is found that if an interchange of two wave numbers makes the only difference between two terms, then the ratio of the corresponding coefficients may be written in terms of precisely the same phase shift $\phi(k_i, k_j)$ as that which appeared in the wavefunction with two reversed spins. For example,

$$\frac{A_{213}}{A_{123}} = \mathrm{e}^{\mathrm{i}\phi(k_2, k_1)} \,. \tag{15.5.88}$$

When two wave number permutations are needed to reach one term from another, the amplitude ratio contains the sum of the two phase shifts associated with the permutations. For example,

[10] H. BETHE, 1931. HANS ALBRECHT BETHE (1906–2005) was awarded the Nobel Prize in 1967 "for his contributions to the theory of nuclear reactions, especially his discoveries concerning the energy production in stars".

$$\frac{A_{312}}{A_{123}} = \mathrm{e}^{\mathrm{i}(\phi(k_3,k_1)+\phi(k_3,k_2))} \,. \tag{15.5.89}$$

This result may be interpreted as follows. The scattering of three magnons by each other can be understood in terms of a series of two-particle scattering events. The phase shifts in each two-particle process are of the previously seen form, irrespective of the state of the third particle. In the language of scattering theory: the S-matrix of three-particle scattering is factorizable.

Furthermore the boundary condition leads to the generalization of (15.5.76),

$$\begin{aligned} \mathrm{e}^{\mathrm{i}k_1 Na} &= \mathrm{e}^{\mathrm{i}(\phi(k_1,k_2)+\phi(k_1,k_3))} \,, \\ \mathrm{e}^{\mathrm{i}k_2 Na} &= \mathrm{e}^{\mathrm{i}(\phi(k_2,k_1)+\phi(k_2,k_3))} \,, \\ \mathrm{e}^{\mathrm{i}k_3 Na} &= \mathrm{e}^{\mathrm{i}(\phi(k_3,k_1)+\phi(k_3,k_2))} \,, \end{aligned} \tag{15.5.90}$$

or equivalently, to the generalization of (15.5.77),

$$\begin{aligned} k_1 Na &= 2\pi I_1 + \phi(k_1,k_2) + \phi(k_1,k_3) \,, \\ k_2 Na &= 2\pi I_2 + \phi(k_2,k_1) + \phi(k_2,k_3) \,, \\ k_3 Na &= 2\pi I_3 + \phi(k_3,k_1) + \phi(k_3,k_2) \,, \end{aligned} \tag{15.5.91}$$

where I_1, I_2 and I_3 are integers such that $0 \le I_i < N$.

The spectrum of the allowed energies $\hbar\omega(k)$ with $k = k_1 + k_2 + k_3$ has a broad continuum corresponding to the scattering states of three essentially independent magnons, but in addition two-magnon bound states and also three-magnon bound states appear in it. The dispersion relation for the latter is

$$\hbar\omega = J\left[\Delta - \frac{2\Delta + \cos ka}{4\Delta^2 - 1}\right] \,. \tag{15.5.92}$$

In the isotropic ferromagnetic point the energy expression of this bound state simplifies to

$$\hbar\omega = \tfrac{1}{3} J\left[1 - \cos ka\right] \,. \tag{15.5.93}$$

The energy of the three-magnon bound state is therefore lower than that of its two-magnon counterpart.

In m-magnon states, that is when m spins are reversed, the wavefunction may be written as

$$|\Psi_m\rangle = \sum_{x_{l_1} < x_{l_2} < \cdots < x_{l_m}} c(x_{l_1}, \ldots x_{l_m}) \big|\psi(x_{l_1}, \ldots, x_{l_m})\big\rangle \,, \tag{15.5.94}$$

where

$$\big|\psi(x_{l_1}, \ldots, x_{l_m})\big\rangle = \sigma^-_{l_1} \ldots \sigma^-_{l_m} |F\rangle \,. \tag{15.5.95}$$

According to the Bethe ansatz, this state may be characterized by m different wave numbers $k_1, \ldots, k_m$,

$$c(x_{l_1}, \dots x_{l_m}) = \sum_P A(P) \exp \mathrm{i} \left[\sum_{j=1}^{m} k_{P_j} x_{l_j} \right], \tag{15.5.96}$$

where P stands for all possible permutations of the wave numbers. The excitation energy and total wave number of the state appear again as the sum for m apparently independent magnons,

$$\Delta E = J \sum_{j=1}^{m} (\Delta - \cos k_j a) \qquad \text{and} \qquad k = \sum_{j=1}^{m} k_j \,. \tag{15.5.97}$$

In fact, the wave numbers k_j cannot be chosen arbitrarily. Once more, the phase factors $\phi(k_i, k_j)$ are involved in the coefficients $A(P)$ of the wavefunction, therefore, when periodic boundary conditions are used to determine the wave numbers, we get

$$k_i N a = 2\pi I_i + \sum_j \phi(k_i, k_j) \,, \tag{15.5.98}$$

where I_i is again an integer, and the phase shift can be calculated from the generalization of (15.5.75),

$$\cot \tfrac{1}{2} \phi(k_i, k_j) = \frac{\Delta \sin[(k_i - k_j)a/2]}{\cos[(k_i + k_j)a/2] - \Delta \cos[(k_i - k_j)a/2]} \,. \tag{15.5.99}$$

The solutions of these equations provide a wide spectrum of excitations anew, and bound states appear as well. The excitation energy of the lowest-lying bound state is

$$\Delta E = \frac{J}{m} (1 - \cos ka) \,. \tag{15.5.100}$$

15.5.4 The Ground State of the Antiferromagnetic Chain

It follows from our earlier considerations on the anisotropic Heisenberg model that the fully aligned ferromagnetic state can be the true ground state for $\Delta \geq 1$ only. Magnons, the continuum of multi-magnon excitations, and their bound states appear as low-energy excitations above the energy $E_0 = -\frac{1}{4} N J \Delta$ of the ground state in this case. When $\Delta < -1$, the ground state exhibits antiferromagnetic order, with the sublattice magnetization pointing in the direction of the z-axis. Between the two, in the planar regime $-1 < \Delta < 1$, the spin components in the (x, y) plane are more strongly coupled than the component along the quantization axis, hence in the classical limit the spins would lie in the (x, y) plane, exhibiting ferromagnetic order, and the mean value of the z component would vanish. However, in the ground state of one-dimensional models no continuous symmetry of the Hamiltonian can be broken unless the order parameter is conserved. This means that in the anisotropic Heisenberg

model the continuous rotational symmetry around the z-axis cannot be broken in the region $-1 < \Delta < 1$. In addition to the z component, the mean values of the x and y components must also vanish. In the ground state half of the spins must point upward and the other half downward, without any spatial or temporal regularity in the spin fluctuation pattern. This ground state can be found using the Bethe ansatz in the subspace where $m = N/2$ spins are reversed with respect to the fully aligned state. The situation is the same in the ground state of the isotropic antiferromagnetic chain. We shall first study this case.

As it was mentioned in connection with (15.5.57), the isotropic antiferromagnetic model can be defined in two ways. In the customary approach the exchange interaction is assumed to be negative, while in the alternative one adopted there $J > 0$ and $\Delta = -1$ are assumed. The results obtained for the ferromagnetic system can be more easily utilized in the first approach: the equations derived from the Bethe ansatz for the wave numbers and phase shifts are identical to those of the isotropic ferromagnetic model. Therefore we shall use the Hamiltonian

$$\mathcal{H} = J \sum_l \left[\tfrac{1}{2} \left(\sigma_l^+ \sigma_{l+1}^- + \sigma_l^- \sigma_{l+1}^+ \right) + \tfrac{1}{4} \left(\sigma_l^z \sigma_{l+1}^z - 1 \right) \right] + \tfrac{1}{4} NJ \qquad (15.5.101)$$

with $J > 0$ to study the isotropic antiferromagnetic chain.

Since $N/2$ spins are reversed, the same number of different wave numbers have to appear in the Bethe ansatz of the wavefunction, and they have to satisfy the following system of equations:

$$Nak_i = 2\pi I_i + 2 \sum_{j \neq i} \operatorname{arccot} \left[\tfrac{1}{2} \left(\cot(k_i a/2) - \cot(k_j a/2) \right) \right] . \qquad (15.5.102)$$

The energy and total wave number of the corresponding state are

$$E = \tfrac{1}{4} NJ - J \sum_i (1 - \cos k_i a) \qquad \text{and} \qquad k = \sum_i k_i \, . \qquad (15.5.103)$$

In terms of the rapidities $\lambda_i = \cot(k_i a/2)$ instead of the wave number k_i,

$$2N \operatorname{arccot} \lambda_i = 2\pi I_i + 2 \sum_{j \neq i} \operatorname{arccot} \left(\frac{\lambda_i - \lambda_j}{2} \right) , \qquad (15.5.104)$$

which can be transformed into the algebraic form

$$\left(\frac{\lambda_i + \mathrm{i}}{\lambda_i - \mathrm{i}} \right)^N = \prod_{j \neq i} \left(\frac{\lambda_i - \lambda_j + 2\mathrm{i}}{\lambda_i - \lambda_j - 2\mathrm{i}} \right) , \qquad (15.5.105)$$

while the energy takes the form

$$E = \tfrac{1}{4} NJ - J \sum_i \frac{2}{1 + \lambda_i^2} \, . \qquad (15.5.106)$$

The solution of the Bethe equations would lead us too far afield. We shall content ourselves with noting that when half of the spins are reversed ($N_\downarrow = N/2$), there is only one state in which all wave numbers and rapidities are real and finite. In this state, which turns out to be the ground state, the quantum numbers I_i take all odd integer values in the interval $1 \leq I_i < N$. Provided that N is an integral multiple of four, the wave number of the ground state is $k = 0$ and the state is a singlet, i.e., its total spin is zero.

The numerical solution of the Bethe equations reveals that the wave numbers fill the whole interval $(0, 2\pi/a)$, albeit not uniformly. Instead of $k_i = 2\pi I_i/Na$ they tend to be located more densely around the middle of the interval on account of the phase shifts arising from scattering processes, as shown in Fig. 15.7.

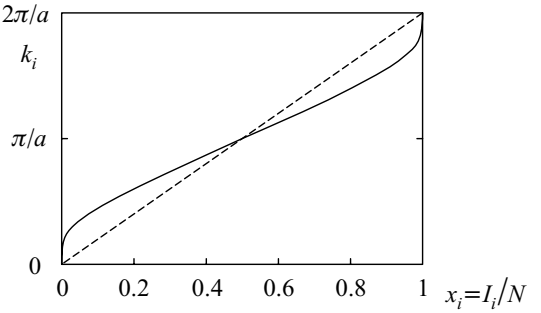


Fig. 15.7. The wave numbers k_i obtained from the numerical solution of the Bethe ansatz equations for the ground state of the isotropic antiferromagnetic Heisenberg model. The dashed line shows the wave numbers for the case when the phase shift is neglected

The ground-state energy can be calculated exactly in the large-N limit:

$$E_0 = \tfrac{1}{4} NJ - NJ \ln 2 = -0.443\, NJ\,. \tag{15.5.107}$$

This energy is lower than that of the Néel state, $E_0 = -\frac{1}{4} NJ$, which would be obtained if there were no phase shifts.

Before proceeding any further it is useful to give the ground state in another parametrization. If instead of the previously used formula $\lambda_i = \cot(k_i a/2)$ rapidity is defined by

$$\lambda_i = -\cot(k_i a/2)\,, \tag{15.5.108}$$

then the following equations are obtained instead of (15.5.104):

$$2N \arctan \lambda_i = 2\pi J_i + 2 \sum_{j \neq i} \arctan\left(\frac{\lambda_i - \lambda_j}{2}\right)\,. \tag{15.5.109}$$

Provided the chain consists of an even number of atoms, J_i is an integer if the number of reversed spins is odd, and a half-integer if it is even. The energy and wave number of the state are given by

$$E = \tfrac{1}{4}NJ - J\sum_i \frac{2}{1+\lambda_i^2} \tag{15.5.110}$$

and

$$k = N_\downarrow \frac{\pi}{a} + \frac{2\pi}{Na}\sum_{i=1}^{N_\downarrow} J_i\,. \tag{15.5.111}$$

Using this parametrization all rapidities and wave numbers are real if each J_i is in the interval

$$-\tfrac{1}{2}(N-N_\downarrow) < J_i < \tfrac{1}{2}(N-N_\downarrow)\,. \tag{15.5.112}$$

When $N_\downarrow = N/2$, there are exactly $N/2$ consecutive integers or half-integers in the above interval:

$$-\tfrac{1}{4}N + \tfrac{1}{2}, -\tfrac{1}{4}N + \tfrac{1}{2} + 1, \ldots, \tfrac{1}{4}N - \tfrac{1}{2}\,. \tag{15.5.113}$$

When this set is chosen for the quantum numbers J_i, $N/2$ real rapidities are obtained. In this parametrization this particular choice of the Bethe quantum numbers and the corresponding rapidities yield the ground state.

15.5.5 Spinon Excitations in the Antiferromagnetic Chain

The excited states above the singlet ground state can be characterized on the one hand their wave number, and on the other hand by their total spin and its z component. The spin of low-energy excitations is expected to differ from the ground-state spin by at most one unit, therefore excited states will be sought in the singlet and triplet subspaces. The distribution of the Bethe quantum numbers I_i, or J_i, will be different, too. For low-energy excitations the set of these quantum numbers is expected to differ little from the corresponding set of the ground state. In what follows, the quantum numbers J_i will be used.

We shall first consider triplet excitations with a spin projection $S^z_{\rm tot} = 1$. Since in these states the number of reversed spins is one less than in the ground state, $N_\downarrow = N/2 - 1$, specifying this state requires one less quantum number k_i or J_i, too. On the other hand, the Bethe equations allow $N - N_\downarrow = N/2 + 1$ different values for J_i. Since the parity of the number of reversed spins is changed with respect to the ground state, the quantum numbers J_i have to be chosen from the sequence

$$-\tfrac{1}{4}N\,, -\tfrac{1}{4}N + 1\,, \ldots\,, \tfrac{1}{4}N - 1, \tfrac{1}{4}N\,. \tag{15.5.114}$$

Real wave numbers are obtained if $N/2-1$ different numbers are chosen out of the $N/2+1$ possible values for J_i. In other words, out of the $N/2+1$ allowed

Bethe quantum numbers $N/2-1$ are occupied and two are empty in this state. Thus a set of states characterized by two parameters (the position of the two holes) is obtained. The energies of these excited states do not determine a sharp dispersion curve but a continuum whose lower and upper bounds are given by

$$\hbar\omega_{\min} = \tfrac{1}{2}\pi J|\sin ka|\,, \qquad \hbar\omega_{\max} = \pi J|\sin(ka/2)|\,. \tag{15.5.115}$$

Similar results are obtained for the $S^z_{\rm tot} = 0$ component of the triplet. This continuum of excitations is shown in Fig. 15.8.

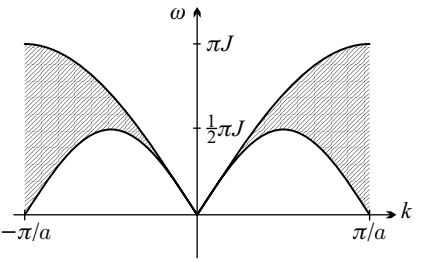


Fig. 15.8. The continuum of triplet excitations in an isotropic antiferromagnetic spin-1/2 chain

Excitation energies can be determined analytically in the $N \to \infty$ limit. After some rather lengthy calculation the energy and wave number of these excitations can be written in the form

$$E - E_0 = \tfrac{1}{2}\pi J \sin k_1 a + \tfrac{1}{2}\pi J \sin k_2 a\,, \qquad k = k_1 + k_2\,, \tag{15.5.116}$$

where k_1 and k_2 lie in the interval $(0, \pi/a)$. Each true (physically realisable) triplet excitation appears to be composed of a pair of "elementary" excitations with the dispersion relation

$$\varepsilon_{\rm s}(k) = \tfrac{1}{2}\pi J \sin ka\,. \tag{15.5.117}$$

This picture becomes even more pronounced when singlet excited states are considered. States are then characterized by $N/2$ quantum numbers, but some of the J_i are identical, and therefore there are complex conjugate pairs among the wave numbers and rapidities. The low-energy part of the excitation spectrum can be determined analytically in the large-N limit, and exactly the same result is obtained as for triplet excitations – i.e., singlet excitations can also be interpreted as pairs of "elementary" excitations whose dispersion relation is given by (15.5.117). We may therefore say that the elementary excitations of the antiferromagnetic Heisenberg chain are spin-1/2 *spinons*, defined over only half of the Brillouin zone, in the region $(0, \pi/a)$, but in any physical process they are created in pairs. The physical meaning of this

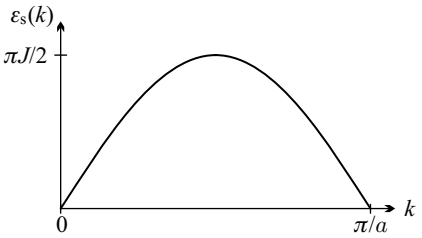


Fig. 15.9. Dispersion relation for spinons in an isotropic antiferromagnetic chain

statement will become clear soon. The dispersion relation of the spinons is shown in Fig. 15.9.

Freely moving spinon excitations can be best illustrated by considering a chain in which the spins are arranged in a Néel-type antiferromagnetic order, except at one site, where the spin is reversed. Compared to the antiferromagnetic sequence, two bonds are not satisfied. Since this state is not an eigenstate of the Heisenberg Hamiltonian, two oppositely directed spins can flip each other. When one of the spins is in a "bad" bond, the antiferromagnetic order will be reestablished in this bond after the spin flip but, as shown in Fig. 15.10, a next-nearest bond will become unsatisfied, with two parallel spins. As the spin flip processes continue, the two "bad" bonds move independently. A Néel-type antiferromagnetic order exists between them, but in the opposite phase. "Bad" bonds can therefore be considered as "domain walls". In this picture spinons are these moving domain walls. Their motion makes the Néel-type order unstable, and gives rise to a spin-liquid state.

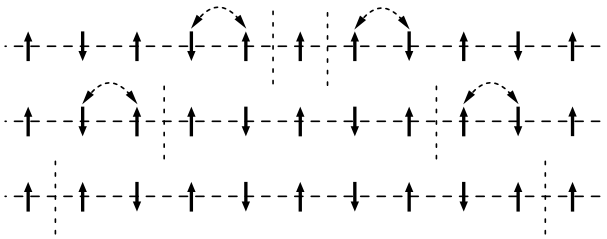


Fig. 15.10. Freely propagating spinons (domain walls) in an antiferromagnetic Heisenberg chain with one reversed spin. Dotted lines denote the „bad" bonds, and the arrows point to the spins that will be flipped in the next step

When two spins next to a domain wall are flipped by the exchange interaction, the domain wall jumps by two lattice constants. Constructing the spinon wavefunction as the linear combination of spin configurations containing such domain walls, states of wave number k and $k + \pi/a$ are equivalent since the spinon propagates along the chain as if the lattice constant were doubled. This

explains why the dispersion relation is defined over only half of the Brillouin zone, in the range $(0, \pi/a)$.

When the spectrum of excited states and the corresponding wavefunctions are known, the correlation function between two spins a distance r apart could be calculated, in principle. However, the calculation cannot be performed analytically along these lines. We shall return to this problem in the third volume. Here we just mention that the correlation functions shows a power-law decay, as expected for a gapless critical model.

15.5.6 The One-Dimensional XY Model

Since the Bethe ansatz is satisfied even in the anisotropic spin-1/2 Heisenberg model, as has been discussed for the anisotropic ferromagnetic case, the behavior of the system can be studied in the planar regime, $-1 < \Delta < 1$, too. The behavior is qualitatively similar to that observed in the isotropic antiferromagnetic point. The ground state is a singlet, low-lying singlet and triplet excitations form a gapless continuum, and the energy and wave number of the excitations inside the continuum can be constructed as if they consisted of pairs of spin-1/2 spinons with dispersion relation

$$\varepsilon_{\mathrm{s}}(k) = \tfrac{1}{2}\pi J \frac{\sin\Theta}{\Theta} \sin ka \,, \tag{15.5.118}$$

where Θ is related to the anisotropy parameter Δ via $\cos\Theta = \Delta$,[11] and the spinons are defined in the interval $(0, \pi/a)$. The continuum of excitations is similar to that shown in Fig. 15.8 for the isotropic antiferromagnetic point, only the scale is different. The boundaries of the continuum are given by

$$\hbar\omega_{\min} = \tfrac{1}{2}\pi J \frac{\sin\Theta}{\Theta} |\sin ka| \,, \qquad \hbar\omega_{\max} = \pi J \frac{\sin\Theta}{\Theta} |\sin(ka/2)| \,. \tag{15.5.119}$$

To get an even better picture of why low-energy excitations appear as pairs of fictitious particles, it is worthwile to study the special case $\Delta = 0$. This is the XY model. The Jordan–Wigner transformation (15.2.62) allows us to express the Hamiltonian of this model in terms of spinless fermions in a particularly simple form:

$$\mathcal{H} = -\tfrac{1}{2} J \sum_j \left(c_j^\dagger c_{j+1} + c_{j+1}^\dagger c_j \right) . \tag{15.5.120}$$

The negative sign is chosen for later convenience, and $J > 0$ is assumed. Using the Fourier transforms of the creation and annihilation operators, the Hamiltonian is rewritten as

[11] Here, too, the choice $J > 0$, $\Delta = 1$ corresponds to the isotropic antiferromagnetic point.

$$\mathcal{H} = \sum_k \varepsilon_k c_k^\dagger c_k \,, \tag{15.5.121}$$

where $\varepsilon_k = -J \cos ka$. The magnetic system is thus equivalent to a gas of free spinless fermions with a simple cosine dispersion relation. To specify the allowed values of the wave number it should be noted that (15.5.99) implies that for $\Delta = 0$ the phase shifts $\phi(k_i, k_j)$ assume the values $\pm\pi$. Depending on the number of reversed spins relative to the ferromagnetic ground state with all spins pointing upward, the net phase shift is either 0 or π. Denoting the number of downward spins by $N_\downarrow$,

$$k_i = \frac{2\pi}{Na}(I_i + \tfrac{1}{2}) \,, \tag{15.5.122}$$

if $N_\downarrow$ is even, and

$$k_i = \frac{2\pi}{Na} I_i \,, \tag{15.5.123}$$

if $N_\downarrow$ is odd. Figure 15.11 shows the spectrum of a finite chain with the possible values of k_i.

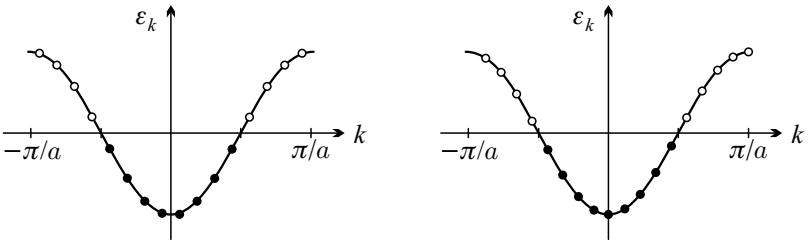


Fig. 15.11. The spectrum of free spinless fermions in the XY model, and the allowed values of the wave number for $N_\downarrow$ even and odd, respectively. $N = 16$, $N_\downarrow = 8$ on the left side, and $N = 18$, $N_\downarrow = 9$ on the right side. Full (empty) circles indicate states that are occupied (unoccupied) in the ground state

When $J > 0$, wave numbers in the range $-\pi/2a \le k \le \pi/2a$ appear in the ground-state wavefunction, since the associated one-particle states have negative energy. One may say that states characterized by such wave numbers are occupied. As k has as many possible values in the Brillouin zone $-\pi/a \le k \le \pi/a$ as there are lattice sites, and half of the corresponding states are occupied, one may speak of a half-filled band. The Jordan–Wigner transformation also implies that

$$\langle S_j^z \rangle = \langle c_j^\dagger c_j \rangle - \tfrac{1}{2} = \frac{1}{N} \sum_k \langle c_k^\dagger c_k \rangle - \tfrac{1}{2} = 0 \tag{15.5.124}$$

in this half-filled case, that is the expectation values of the spin and the magnetic moment vanish at each lattice site in the ground state. Likewise,

the expectation values for the x and y components vanish, too. Furthermore, if the number of lattice sites is even, then the magnitude of the total spin also vanishes – in other words, the system has a singlet ground state. The spins are completely disordered, therefore we may call the ground state a *spin liquid*.

There are two ways to create low-energy excitations above this ground state. The first option is leaving the number of spinless fermions unaltered, and choosing the wave numbers in the wave function differently – in other words, changing the distribution of occupied states. In the simplest case instead of a state of wave number k_i filled in the ground state, another state of wave number k_j will be occupied. Wave numbers associated with such low-energy excited states are shown in Fig. 15.12.

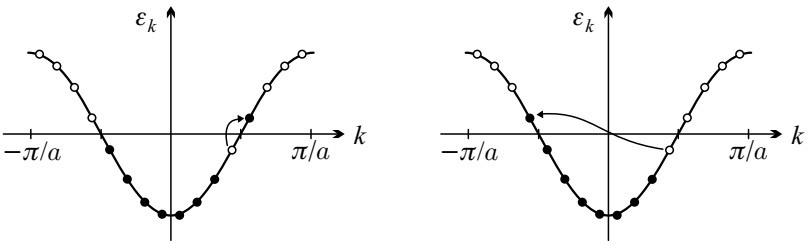


Fig. 15.12. Two possible wave-number configurations for $N = 16$ and $N_\downarrow = 8$ in low-energy excited states in the XY model. Occupied states are indicated by full circles

In terms of spinless fermions, these excitations of the spin model may be interpreted as particle–hole excitations. Taking away a particle with wave number k_1 and adding another with wave number k_2, the energies of excitations with wave numbers $k = k_2 - k_1$ form a continuum, as depicted in Fig. 15.13. The continuum is bounded by

$$\hbar\omega_{\min} = J|\sin ka| \qquad \text{and} \qquad \hbar\omega_{\max} = 2J|\sin(ka/2)|\,. \tag{15.5.125}$$

This result was derived in light of the fact that the creation of holes is possible only in the interval $-\pi/2a < k_1 < \pi/2a$, while that of particles only in the complementary range of the Brillouin zone.

Another type of excitation is obtained when a spinless fermion is added to or removed from the system – corresponding to raising or lowering the z component of the total spin by unity. In contrast to the two-particle excitations discussed above, they might seem to be one-particle excitations. In reality, they are particle–hole excitations, too, but with respect to a modified configuration. Namely, it has to be taken into account that spin reversal changes the parity of $N_\downarrow$ with respect to the ground state, and therefore all wave numbers are shifted. In the lowest-energy state of the subspace $S^z_{\text{tot}} = 1$ the wave numbers are located symmetrically anew. To obtain excited states above it,

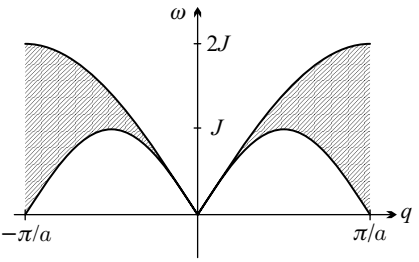


Fig. 15.13. The continuum of low-energy excitations in the XY model

a particle–hole pair has to be created in this distribution. The distribution of the occupied states is shown in Fig. 15.14.

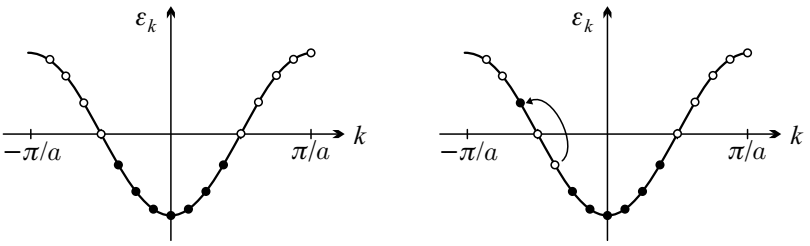


Fig. 15.14. Wave numbers of free spinless fermions in the lowest-energy state of the $S^z_{\rm tot} = 1$ sector and in a possible excited configuration for $N = 16$ and $N_\downarrow = 7$. Occupied sites are indicated by full circles

It can be shown that in the $N \to \infty$ limit the continuum of these excitations coincides with the continuum arising in the subspace $S^z_{\rm tot} = 0$.

This mapping to free spinless fermions is valid for $\Delta = 0$ only. The Jordan–Wigner transformation can be applied to the case when Δ is finite, but then an interaction appears among the spinless fermions. We shall defer the discussion of the effects of this interaction to the third volume, where we shall derive the above-mentioned result – namely that the picture presented for the isotropic antiferromagnet and the XY model is qualitatively true in the whole planar regime $-1 < \Delta < 1$, where the spin projection along the z direction is expected to vanish classically. In fact, the quantum mechanical ground state is in the subspace $S^z_{\rm tot} = 0$ and the continuum of low-energy excitations can be interpreted as arising from pairs of spin-1/2 spinons.

15.5.7 The Role of Next-Nearest-Neighbor Interactions

So far it has been assumed that the antiferromagnetic exchange interaction acts between nearest neighbor spins only. A different type of disordered spin

configuration can be found when the interaction between farther neighbors is not negligible, since in addition to quantum fluctuations spin frustration that may arise from the competition between magnetic couplings has to be taken into account, too.

We shall denote the strength of the antiferromagnetic coupling between nearest neighbors by J_1, and between next-nearest neighbors by J_2. The Hamiltonian of the spin chain is then

$$\mathcal{H} = J_1 \sum_i \boldsymbol{S}_i \cdot \boldsymbol{S}_{i+1} + J_2 \sum_i \boldsymbol{S}_i \cdot \boldsymbol{S}_{i+2} \,. \tag{15.5.126}$$

Both J_1 and J_2 are assumed to be positive. Spin frustration is best illustrated on a zigzag ladder, as shown in Fig. 15.15. Spins located at even and odd sites of the original chain form two antiferromagnetic chains – the legs of the ladder – that are somewhat displaced relative to each other. Each spin on one leg is coupled antiferromagnetically to two spins on the other leg.

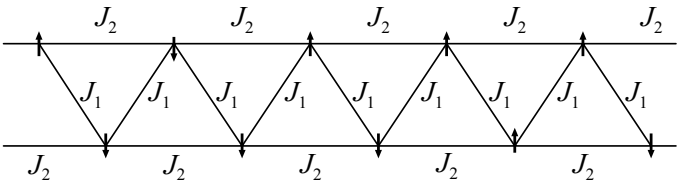


Fig. 15.15. Spin chain with nearest- and next-nearest-neighbor interactions illustrated by a zigzag ladder

Separating the terms within the legs and between them, the system is described by the Hamiltonian

$$\begin{aligned} \mathcal{H} = {} & J_1 \sum_i \left[\boldsymbol{S}_{2i-1} \cdot \boldsymbol{S}_{2i} + \boldsymbol{S}_{2i} \cdot \boldsymbol{S}_{2i+1} \right] \\ & + J_2 \sum_i \left[\boldsymbol{S}_{2i-1} \cdot \boldsymbol{S}_{2i+1} + \boldsymbol{S}_{2i} \cdot \boldsymbol{S}_{2i+2} \right] . \end{aligned} \tag{15.5.127}$$

The coupling J_2 tends to align the spins on the legs antiferromagnetically – at least on short distances. Assuming that two neighboring spins on a leg are aligned oppositely (as they should be if the antiferromagnetic coupling between them is strong enough), they create opposite effective fields at the position of the spin lying between them on the other leg, canceling each other's effect. If, on the other hand, the coupling J_1 is stronger, then a short-range antiferromanetic order may be expected along the zigzag path, which would give rise to a parallel alignment of neighboring spins on the same leg. When J_1 and J_2 are of comparable strength, spins become frustrated, and a new type of singlet phase appears.

The ground state of this model is exactly known in the Majumdar–Ghosh point,[12] where $J_2 = \frac{1}{2}J_1$. In this special case the Hamiltonian can be written as

$$\mathcal{H} = \tfrac{3}{4}J_1 \sum_i \left[P_{3/2}(\boldsymbol{S}_{2i-1}+\boldsymbol{S}_{2i}+\boldsymbol{S}_{2i+1})+P_{3/2}(\boldsymbol{S}_{2i}+\boldsymbol{S}_{2i+1}+\boldsymbol{S}_{2i+2})\right]-3J_1N/8\,, \tag{15.5.128}$$

where the operator

$$P_{3/2}(\boldsymbol{S}_{2i-1}+\boldsymbol{S}_{2i}+\boldsymbol{S}_{2i+1}) = \tfrac{1}{3}\Big[(\boldsymbol{S}_{2i-1}+\boldsymbol{S}_{2i}+\boldsymbol{S}_{2i+1})^2-\tfrac{1}{2}\big(\tfrac{1}{2}+1\big)\Big] \tag{15.5.129}$$

projects to the subspace where the three $S = 1/2$ spins add up to a total spin of $3/2$. It is immediately seen from this form that the state in which singlet pairs are formed between nearest neighbors along the zigzag path is an exact ground state as three consecutive spins can never be combined into a total spin of $3/2$. As shown in Fig. 15.16, there are two ways to form singlet pairs, so the system has a doubly degenerate ground state. Since the singlet pairs are independent of each other there is no correlation between the spins beyond the nearest-neighbor distance.

Fig. 15.16. The two equivalent ground states of the antiferromagnetic spin chain in the zigzag ladder representation for $J_2 = \frac{1}{2}J_1$. The spins connected by solid lines form singlet pairs

Breaking any of the singlet pairs gives rise to an excited state. A relatively low-energy excitation is expected when the two spins of the broken singlet are not bound into a triplet but move farther away, and the singlet bonds between them are rearranged, as shown in Fig. 15.17. The configuration of singlet bonds corresponds to different ground states at the two ends and in the intermediate region. Since the energy of the singlet pairs is not modified by the rearrangement, the energy of the configuration is independent of the length of the rearranged region, i.e., on the distance between the free spins. Therefore they can propagate freely. Consequently the continuum of low-energy excitations can be interpreted as arising from a pair of spinons or domain walls, the creation of which requires a finite amount of energy that is equal to the binding energy of the singlets.

Such a behavior is observed not only in the point $J_2 = \frac{1}{2}J_1$ but whenever the antiferromagnetic second-neighbor coupling is stronger than a critical value $(J_2/J_1)_c = 0.241$.

[12] C. K. Majumdar and D. K. Ghosh, 1969.

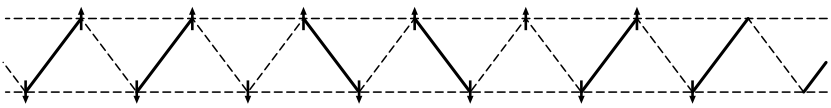


Fig. 15.17. Freely moving spinons in the spin-1/2 antiferromagnetic zigzag ladder

In the previous discussion we have seen two possible types of behavior for the $S = 1/2$ spin chains. The ground state was either a unique singlet state with a gapless excitation spectrum, or the ground state was doubly degenerate, dimerized, and the excitation spectrum was separated from the ground state by a finite gap. According to the *Lieb–Schultz–Mattis theorem*,[13] these two types of behavior are generic in one-dimensional spin models with half-odd-integer spin. More precisely, the authors established that in half-odd-integer spin models that are invariant under translation by the lattice constant and have full rotational symmetry in spin space, the ground state and the excitation continuum may be separated by a gap only if the ground state is degenerate and breaks the translational symmetry spontaneously.

15.5.8 Excitations in the Spin-One Heisenberg Chain

In contrast to the spin-1/2 chain, the ground state and the excitation spectrum of antiferromagnetic Heisenberg models with $S > 1/2$ cannot be determined exactly. States with more than two reversed spins cannot be given in the Bethe ansatz form, and scattering processes involving several magnons cannot be described in terms of the phase shifts occurring in two-magnon scattering. To obtain exactly solvable models, terms involving higher powers of the product of two spins, or the products of several spins – with suitably chosen coupling constants – must be included in the Hamiltonian. In the $S = 1$ case, the model defined by the Hamiltonian

$$\mathcal{H} = -J \sum_i \left[\boldsymbol{S}_i \cdot \boldsymbol{S}_{i+1} - (\boldsymbol{S}_i \cdot \boldsymbol{S}_{i+1})^2 \right], \tag{15.5.130}$$

while in the $S = 3/2$ case, the system with the Hamiltonian

$$\mathcal{H} = -J \sum_i \left[\boldsymbol{S}_i \cdot \boldsymbol{S}_{i+1} - \tfrac{8}{27} (\boldsymbol{S}_i \cdot \boldsymbol{S}_{i+1})^2 - \tfrac{16}{27} (\boldsymbol{S}_i \cdot \boldsymbol{S}_{i+1})^3 \right] \tag{15.5.131}$$

are exactly solvable by straightforward generalization of the Bethe ansatz. Both the ground state and the excitation spectrum can be determined, and the behavior of the system is found to be similar to that of the spin-1/2 isotropic Heisenberg model. When the coupling is ferromagnetic, the ground state is fully aligned, and low-lying excitations are spin-1 magnons. Antiferromagnetic coupling, on the other hand, leads to a singlet ground state with a gapless

[13] E. Lieb, T. D. Schultz, and D. C. Mattis, 1962.

continuum of excitations that consists of singlet and triplet pairs of spin-1/2 spinons.

Another exactly solvable version of the spin-1 model is obtained by reversing the sign of the biquadratic term:

$$\mathcal{H} = -J \sum_i \left[\boldsymbol{S}_i \cdot \boldsymbol{S}_{i+1} + (\boldsymbol{S}_i \cdot \boldsymbol{S}_{i+1})^2 \right] . \tag{15.5.132}$$

This model possesses not only SU(2) but also SU(3) symmetry. Here, too, the continuum of excitations appears above the ground state without gap, however, the soft modes appear at $k = 0$ and $\pm 2\pi/3a$ in this model, and not at $k = 0$ and π/a as in the spin-1/2 antiferromagnetic Heisenberg chain.

These models should nevertheless be considered as exceptions. In general, isotropic magnetic chains with integer spin do not have a gapless branch of excitations. The Lieb–Schultz–Mattis theorem is not applicable to these models, and, barring certain exceptions, even if the ground state is nondegenerate, the continuum of excited states in integer-spin antiferromagnetic chains is separated from the ground state by a forbidden region, the *Haldane gap*.[14]

To understand the nature of this state better, consider the most general Hamiltonian describing an interacting $S = 1$ chain with an SU(2)-invariant isotropic nearest-neighbor interaction. It follows from the properties of spin-1 operators discussed in Appendix F that, besides the usual exchange interaction, a biquadratic term $(\boldsymbol{S}_i \cdot \boldsymbol{S}_{i+1})^2$ also appears in the Hamiltonian, but no higher powers. The most general form satisfying these requirements is

$$\mathcal{H} = J \sum_i \left[\boldsymbol{S}_i \cdot \boldsymbol{S}_{i+1} + \beta \, (\boldsymbol{S}_i \cdot \boldsymbol{S}_{i+1})^2 \right] . \tag{15.5.133}$$

This Hamiltonian can be written in an alternative form by taking into account that the combination of two neighboring $S = 1$ spins gives rise to total spins $S_{\text{pair}} = 0$, 1 and 2. The operators that project onto the respective subspaces are

$$\begin{aligned}
P_0(i,i+1) &= 1 - \tfrac{2}{3}(\boldsymbol{S}_i + \boldsymbol{S}_{i+1})^2 + \tfrac{1}{12}(\boldsymbol{S}_i + \boldsymbol{S}_{i+1})^4 \\
&= -\tfrac{1}{3} + \tfrac{1}{3} \, (\boldsymbol{S}_i \cdot \boldsymbol{S}_{i+1})^2 , \\
P_1(i,i+1) &= \tfrac{3}{4}(\boldsymbol{S}_i + \boldsymbol{S}_{i+1})^2 - \tfrac{1}{8}(\boldsymbol{S}_i + \boldsymbol{S}_{i+1})^4 \\
&= 1 - \tfrac{1}{2} \, (\boldsymbol{S}_i \cdot \boldsymbol{S}_{i+1}) - \tfrac{1}{2} \, (\boldsymbol{S}_i \cdot \boldsymbol{S}_{i+1})^2 , \\
P_2(i,i+1) &= -\tfrac{1}{12}(\boldsymbol{S}_i + \boldsymbol{S}_{i+1})^2 + \tfrac{1}{24}(\boldsymbol{S}_i + \boldsymbol{S}_{i+1})^4 \\
&= \tfrac{1}{3} + \tfrac{1}{2} \, (\boldsymbol{S}_i \cdot \boldsymbol{S}_{i+1}) + \tfrac{1}{6} \, (\boldsymbol{S}_i \cdot \boldsymbol{S}_{i+1})^2 .
\end{aligned} \tag{15.5.134}$$

The most general isotropic model is constructed by taking an arbitrary linear combination,

[14] F. D. M. Haldane, 1983.

$$\mathcal{H} = \sum_{i=1}^{N} \sum_{j=0}^{2S} f_j P_j(i, i+1) \,. \tag{15.5.135}$$

The usual antiferromagnetic Heisenberg model is recovered by choosing $f_j = j(j+1)$.

Let us now consider the special case when only those states contribute to the energy for which every pair of neighboring spins is in the $S_{\text{pair}} = 2$ state. With an appropriately chosen additive constant the Hamiltonian may be written as

$$\mathcal{H} = J \sum_i \left[2P_2(i, i+1) - \tfrac{2}{3} \right] \,, \tag{15.5.136}$$

or, in terms of the spin operators, as

$$\mathcal{H} = J \sum_i \left[\boldsymbol{S}_i \cdot \boldsymbol{S}_{i+1} + \tfrac{1}{3} \left(\boldsymbol{S}_i \cdot \boldsymbol{S}_{i+1} \right)^2 \right] \,. \tag{15.5.137}$$

If we succeed in constructing a state without any pairs of neighbors being in the $S_{\text{pair}} = 2$ state, then the energy of the state in question will simply be $E_0 = -\frac{2}{3} NJ$. Such a state can be put together as follows. First, at each lattice site, the $S = 1$ spin operator is built up of two spin-1/2 operators, σ_i and τ_i, allowing only the symmetric combination. Then, a singlet state is formed between lattice sites i and $i+1$, using one of the 1/2 spins at each lattice site. The other 1/2 spin at lattice site i $(i+1)$ is paired with one of the spins at lattice site $i-1$ $(i+2)$ to give another singlet state. Thus, with the exception of the two end points, every lattice point of the chain is bound to its left and right neighbors by a singlet bond. The bond arrangement is shown schematically in Fig. 15.18.

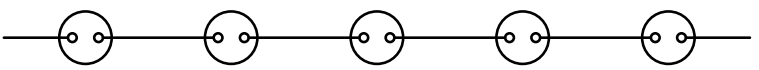


Fig. 15.18. Singlet bonds in the ground state of the model defined by (15.5.137)

When the projection operator P_2 acts on any such configuration, the result is indeed zero. By choosing the linear combination of all possible states with equal weights, symmetrization is taken care of. The resulting state turns out to be more than just an eigenstate of the Hamiltonian (15.5.137): it is the ground state. Since the state comprises singlets between nearest neighbors, and is therefore reminiscent of valence bonds, such models are also called *valence bond solids* (VBS).

To create an excited state one or more short-range valence bonds have to be broken. Since this requires a finite amount of energy, the ground state is separated by a gap from the excitation spectrum in these models. Consequently the spin–spin correlation function decays exponentially, with a correlation length that is not much larger than atomic dimensions. Another interesting point that distinguishes this spin-1 model from the spin-1/2 Heisenberg

model is that the low-lying excitations do not form a continuum. The excitations are not pairs of deconfined spin-1/2 spinons but spin-1 magnons with a well-defined dispersion relation.

If the parameter β in (15.5.133) is not exactly $\beta = 1/3$, then for $-1 < \beta < 1$ (note that this interval contains the isotropic antiferromagnetic Heisenberg model, which corresponds to $\beta = 0$) the ground state and the excitation spectrum are similar to those discussed above – with the exception that singlet bonds are not necessarily formed between neighboring sites, however their average range is short. Once again, the spin–spin correlation functions decay exponentially. This is the Haldane phase of the spin-1 Heisenberg chain.

15.5.9 Spin Ladders

Having overviewed the properties of spin chains, we can now generalize these considerations to two or more coupled spin chains. In the simplest case, where two chains are coupled by an exchange coupling $J_\perp$, the system is described by the Hamiltonian

$$\mathcal{H} = J_\parallel \sum_{i=1}^{N} (\boldsymbol{S}_{1,i} \cdot \boldsymbol{S}_{1,i+1} + \boldsymbol{S}_{2,i} \cdot \boldsymbol{S}_{2,i+1}) + J_\perp \sum_{i=1}^{N} \boldsymbol{S}_{1,i} \cdot \boldsymbol{S}_{2,i} \,. \qquad (15.5.138)$$

When the couplings between the spins are drawn as straight lines, they make up a ladder: $J_\parallel$ is the coupling along the legs, while $J_\perp$ is the coupling between spins on the same rung. This is why the system is called a *spin ladder*. When several chains are coupled, multileg ladders are obtained.

If the rung coupling $J_\perp$ is ferromagnetic and strong enough, the spins on the same rung are bound into a triplet ($S = 1$), and the spin ladder behaves as an $S = 1$ spin chain. When $J_\parallel$ is antiferromagnetic, and periodic boundary conditions are imposed, then the ground state is a nondegenerate singlet state (the analogue of the VBS state discussed above), in which spins on neighboring or not too distant rungs form singlets. Such a configuration is shown in Fig. 15.19(a). Note that if the ladder were cut anywhere between two neighboring rungs an odd number of singlet bonds would be broken.

Figure 15.19(b) shows an excited state in which one singlet bond is broken. Breaking a bond requires a finite amount of energy, thus the excitation spectrum starts with a finite gap, the Haldane gap. It can also be understood from the figure why the two "unbound" spins do not propagate freely, and why the excitation spectrum is not a continuum of spinon pairs (unlike for the zigzag ladder). If this ladder were cut between the two "unbound" spins, an even number of singlet bonds would be broken. The energy of this configuration increases when the two spins are moved farther apart, therefore they do not become truly unbound. Using the language of particle physics, the force increases linearly with distance, and confines the two spinons into a magnon.

A similar, nondegenerate singlet ground state is found when $J_\perp$ is antiferromagnetic: for sufficiently strong coupling spins on the same rung form

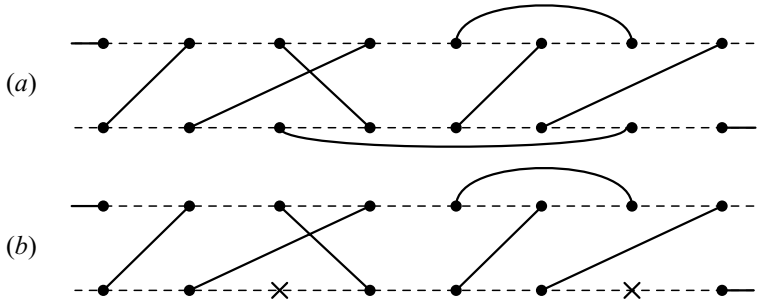


Fig. 15.19. (*a*) A possible configuration of singlet bonds in the ground state of a ferromagnetically coupled two-leg spin ladder. (*b*) An excited state obtained by breaking a singlet bond

singlets with high probability. Such a configuration is shown in Fig. 15.20. Note that in this case the number of singlet bonds between neighboring rungs is always even.

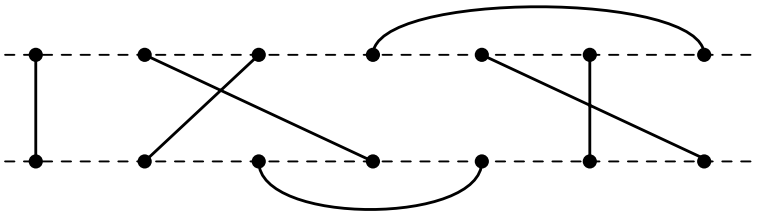


Fig. 15.20. A possible configuration of singlet bonds in the ground state of an antiferromagnetic two-leg spin ladder

The low-lying excitations are magnons again, and the spectrum has a finite energy gap, for the same reason as above: if a singlet bond is broken, the number of singlet bonds between the two "free" spins changes parity, and the energy increases linearly with the distance between the spins, confining the spinons into magnons.

Thus, whether the coupling between the legs is ferromagnetic or antiferromagnetic, two-leg spin ladders behave like spin-1 chains. It can be shown by a generalization of the Lieb–Schultz–Mattis theorem that in odd-leg spin ladders made up of an odd number of $S = 1/2$ spin chains either the ground state is degenerate or the excitation spectrum is gapless. Such spin ladders behave like chains of half-integer spins. On the other hand, the behavior of even-leg spin ladders is analogous to that of chains of integer spins.

15.5.10 Physical Realizations of Spin Chains and Spin Ladders

Ideal, free-standing spin chains and spin ladders (such as the ones studied theoretically in the previous subsections) do not exist in nature. However, there are a number of materials that have truly three-dimensional crystalline structure but quantum mechanical exchange is strongly anisotropic in them: much stronger in one direction than in others. These materials can, therefore, be considered as if they consisted of weakly coupled chains of spins. At low temperatures, where the thermal energy is comparable to or less than the exchange energy between the chains, a three-dimensional ordered magnetic structure may arise. Above this temperature, the chains behave independently as far as magnetic properties are concerned. Here we give just a small selection of materials that exhibit such properties.

A typical example is CPC – $CuCl_2{\cdot}2N(C_5H_5)$ –, in which the chains formed by spin-1/2 Cu^{2+} ions are separated by large pyridine molecules. The exchange interaction between the chains is about 300 times weaker than within the chains, and antiferromagnetic ordering takes place at only $T_N = 1.14\,K$. Copper pyrazine dinitrate (CuPzN) – $Cu(C_4H_4N_2)(NO_3)_2$ – is an even better candidate for the ideal isotropic antiferromagnetic Heisenberg chain. The coupling between the chains is four orders of magnitude weaker than within the chains, and no magnetic ordering has been observed down to 0.1 K. The continuum of low-lying excitations of the spin-1/2 Heisenberg chain, which has been interpreted in terms of spinon pairs, can be observed very well in these materials.

An almost ideal representative of integer-spin antiferromagnetic Heisenberg chains is Y_2BaNiO_5. The spin-1 nickel ions surrounded by oxygens form chains that are extremely weakly coupled. In agreement with theoretical predictions, the excitation spectrum has a finite Haldane gap ($\Delta = 8.6\,meV$).

Several materials have been produced synthetically in which spins form two- or multileg ladders. In the homologous series $Sr_{n-1}Cu_nO_{2n-1}$ spin-1/2 Cu^{2+} ions are arranged in such a way that the copper chains lie in magnetically isolated layers; within the layers n chains form an n-leg ladder, but adjacent ladders are displaced by half a unit, and thus the coupling between ladders is frustrated, and has a negligible effect. Figure 15.21 shows the susceptibility for the two- and three-leg members of this series.

It can be seen that by subtracting the Curie component – which arises from the end spins of the chains or from paramagnetic impurities –, the susceptibility becomes exponentially small at low temperatures for the two-leg ladder, while it remains finite for the three-leg ladder. This is in good agreement with the result explained above: there are no gapless magnetic excitations in the two-leg ladder, while the excitations spectrum of the three-leg ladder is gapless.

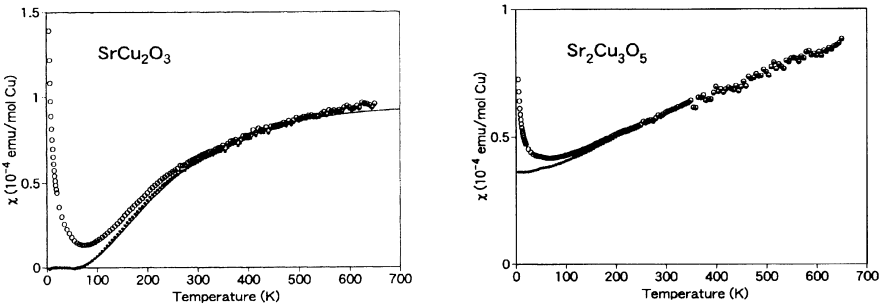


Fig. 15.21. Temperature dependence of the susceptibility of the two-leg ($SrCu_2O_3$) and three-leg ($Sr_2Cu_3O_5$) members of the series. Empty circles indicate raw data, while full circles are corrected data obtained by subtracting the Curie component [M. Azuma et al., *Phys. Rev. Lett.* **73**, 3463 (1994)]

15.6 Spin Liquids

It was shown in the previous sections that quantum fluctuations become relevant in low-dimensional systems. In one dimension they will hinder magnetic ordering, with the sole exception of the ferromagnetic phase. The ground state of antiferromagnetic spin chains and spin ladders is a spin singlet in which the continuous SU(2) symmetry of the Hamiltonian is not broken. Such a state might be called a *spin liquid*.

When considering the excitation spectrum and the decay of the correlation functions in these systems, various fundamentally different types of behavior are found. As we shall see in Chapter 32, owing to the gapless character of the excitation continuum the decay of the correlation function is power-law-like in the spin-1/2 antiferromagnetic Heisenberg model, just like in a critical system. Such systems are called *algebraic* or *critical spin liquids*. When the next-nearest exchange is sufficiently strong and also antiferromagnetic, the competition of the couplings leads to frustration, resulting in an exponentially decaying correlation. Although the situation is similar to the decay of correlations in liquids, the spin system is not homogeneous: the symmetry of the Hamiltonian under translation through the lattice constant is broken. Correlations decay exponentially in the Haldane phase, too, however, a certain order is exhibited by the singlet valence bonds: a hidden order in the spin components.

Thus, none of the above examples is a true spin liquid in the strict sense – namely, that the magnetically disordered singlet ground state should break no symmetry of the Hamiltonian, that there should be no hidden order parameters, and that the spin–spin correlation functions should decay exponentially. They indicate, however, that if a spin liquid phase exists, it may arise from quantum fluctuations, frustration caused by competing interactions, or geometric frustration resulting from the topology of the arrangement of spins.

As a first example consider a two-dimensional square lattice with nearest-neighbor interactions. In this system spins are not frustrated, and quantum fluctuations cannot destroy the Néel-type antiferromagnetic order. However, in the presence of a sufficiently strong ($J_\mathrm{d} \approx 0.5J$) antiferromagnetic coupling between (diagonally separated) next-nearest neighbors the ground state is disordered – although the exact nature of this disorder is not fully understood yet.

Frustration arising from the geometry (topology) of the lattice is observed in the antiferromagnetic triangular lattice. Considering first a single triangle, the three spins at the vertices are each others' nearest neighbors. If, owing to the antiferromagnetic coupling, the spins at two vertices are oriented oppositely (quantum mechanically: if their spin projections are $1/2$ and $-1/2$), then they act oppositely on the third spin. With its orientation undetermined, the state of the third spin is the combination of the $s_z = 1/2$ and $s_z = -1/2$ components with equal weights since the two spin orientations are of the same energy. As a resolution of this frustration ANDERSON[15] suggested that the ground state of the isotropic antiferromagnetic Heisenberg model on a triangular lattice is a disordered singlet. It is a superposition of singlet states in which every spin forms a singlet pair with a nearby spin. In the language of quantum chemistry, the state is a superposition of valence bonds. The energy can be lowered if the singlet pairs break up and reform – that is, if they resonate. Hence the name for this hypothetical spin liquid: *resonating valence bond spin liquid* (RVB spin liquid).

It turned out that resonating singlets do not sufficiently lower the energy on a triangular lattice. The true ground state is a three-sublattice Néel-type state. If, however, in addition to nearest-neighbors interactions other couplings – e.g., multi-spin exchange processes – are also important, the ground state may be a spin-liquid state.

There exist certain lattices in nature that are topologically more strongly frustrated than the triangular lattice, e.g., the two-dimensional kagome lattice or the three-dimensional pyrochlore lattice. The kagome lattice shown in Fig. 5.4 can be described as a network of corner-sharing (interlaced) triangles. In the pyrochlore lattice corner-sharing tetrahedra form a face-centered cubic lattice. Assuming classical spins in both cases, the orientation of the spins on neighboring units is not fixed even when the sum of the spins on each triangle or tetrahedron is required to vanish, and the ground state is disordered. The quantum mechanical ground states of these models are not yet known, but they are expected to be spin liquids. An indication for this is that $Cu_3V_2O_7(OH)_2{\cdot}2H_2O$ – in which the spin-$1/2$ copper ions form a kagome lattice – does not have any usual magnetically ordered state. Similarly, in $Tb_2Ti_2O_7$ – which has a pyrochlore structure, with magnetic rare-earth ions at the vertices of the tetrahedra – no magnetic ordering has been observed down to 70 mK.

[15] P. W. ANDERSON, 1973.

If the RVB state proposed by ANDERSON is realized in these materials, then, on account of the resonance of short-range singlet valence bonds, the two spins liberated in the excited state from the singlet pair are expected to propagate practically independently, and the true elementary excitations are gapped spin-1/2 spinons. According to the numerical results obtained for such models the spectrum of triplet excitations does, indeed, have a finite gap, however, it is filled with a continuum of singlet excitations. This indicates that there are still a lot of mysteries to be resolved about the spin-liquid state.

Further Reading

1. A. Herpin, *Théorie du magnétisme*, Presses Universitaires de France, Paris (1968).
2. L.-P. Lévy, *Magnetism and Superconductivity*, Texts and Monographs in Physics, Springer-Verlag, Berlin (2000).
3. D. C. Mattis, *The Theory of Magnetism I: Thermodynamics and Statistical Mechanics*, Springer-Verlag, Berlin (1988); *The Theory of Magnetism II: Statics and Dynamics*, Springer-Verlag, Berlin (1985).
4. K. Yosida, *Theory of Magnetism*, Springer Series in Solid-State Sciences, Vol. 122, Springer-Verlag, Berlin (1996).
5. S. V. Vonsovskii, *Magnetism*, John Wiley & Sons, New York (1974).
6. R. M. White, *Quantum Theory of Magnetism*, Springer Series in Solid-State Sciences, Vol 32, Springer-Verlag, Berlin (2002).

A

Physical Constants and Units

A.1 Physical Constants

The following table contains the values in SI units for some physical constants that are particularly important in solid-state physics.

Name	Symbol	Value
speed of light		
in vacuum	c	$299\,792\,458\,\mathrm{m\,s^{-1}}$
magnetic constant		
(permeability of free space)	μ_0	$4\pi \times 10^{-7}\,\mathrm{N\,A^{-2}}$
electric constant		
(permittivity of free space)	$\epsilon_0 = 1/\mu_0 c^2$	$8.854\,188 \times 10^{-12}\,\mathrm{F\,m^{-1}}$
elementary charge	e	$1.602\,176 \times 10^{-19}\,\mathrm{C}$
Planck constant	h	$6.626\,069 \times 10^{-34}\,\mathrm{J\,s}$
in eV	$h/\{e\}$	$4.135\,667 \times 10^{-15}\,\mathrm{eV\,s}$
reduced Planck constant	$\hbar = h/2\pi$	$1.054\,572 \times 10^{-34}\,\mathrm{J\,s}$
in eV	$\hbar/\{e\}$	$6.582\,119 \times 10^{-16}\,\mathrm{eV\,s}$
fine-structure constant	$\alpha = e^2/4\pi\epsilon_0\hbar c$	$7.297\,353 \times 10^{-3}$
inverse of α	α^{-1}	137.035 999
magnetic flux quantum	$\Phi_0 = h/2e$	$2.067\,834 \times 10^{-15}\,\mathrm{Wb}$
conductance quantum	$G_0 = 2e^2/h$	$7.748\,092 \times 10^{-5}\,\mathrm{S}$
inverse of G_0	$G_0^{-1} = h/2e^2$	$12\,906.404\,\Omega$
Josephson constant	$K_\mathrm{J} = 2e/h$	$483\,597.9 \times 10^{9}\,\mathrm{Hz\,V^{-1}}$
von Klitzing constant	$R_\mathrm{K} = h/e^2$	$25\,812.807\,\Omega$
Boltzmann constant	k_B	$1.380\,650 \times 10^{-23}\,\mathrm{J\,K^{-1}}$
Avogadro constant	N_A	$6.022\,142 \times 10^{23}\,\mathrm{mol^{-1}}$

Continued on the next page

Name	Symbol	Value
molar gas constant	$R = N_\mathrm{A} k_\mathrm{B}$	$8.314\,472\,\mathrm{J\,mol^{-1}\,K^{-1}}$
Bohr magneton	$\mu_\mathrm{B} = e\hbar/2m_\mathrm{e}$	$9.274\,009 \times 10^{-24}\,\mathrm{J\,T^{-1}}$
nuclear magneton	$\mu_\mathrm{N} = e\hbar/2m_\mathrm{p}$	$5.050\,783 \times 10^{-27}\,\mathrm{J\,T^{-1}}$
Bohr radius	$a_0 = 4\pi\epsilon_0\hbar^2/m_\mathrm{e}e^2$	$0.529\,177 \times 10^{-10}\,\mathrm{m}$
electron mass	m_e	$9.109\,382 \times 10^{-31}\,\mathrm{kg}$
electron magnetic moment	μ_e	$-9.284\,764 \times 10^{-24}\,\mathrm{J\,T^{-1}}$ $-1.001\,160\,\mu_\mathrm{B}$
electron g-factor	$g_\mathrm{e} = 2\mu_\mathrm{e}/\mu_\mathrm{B}$	$-2.002\,319$
electron gyromagnetic ratio	$\gamma_\mathrm{e} = 2\lvert\mu_\mathrm{e}\rvert/\hbar$ $\gamma_\mathrm{e}/2\pi$	$1.760\,860 \times 10^{11}\,\mathrm{s^{-1}\,T^{-1}}$ $28\,024.9532\,\mathrm{MHz\,T^{-1}}$
neutron mass	m_n	$1.674\,927 \times 10^{-27}\,\mathrm{kg}$
neutron magnetic moment	μ_n	$-0.966\,236 \times 10^{-26}\,\mathrm{J\,T^{-1}}$ $-1.913\,043\,\mu_\mathrm{N}$
neutron g-factor	$g_\mathrm{n} = 2\mu_\mathrm{n}/\mu_\mathrm{N}$	$-3.826\,085$
proton mass	m_p	$1.672\,622 \times 10^{-27}\,\mathrm{kg}$
proton–electron mass ratio	$m_\mathrm{p}/m_\mathrm{e}$	$1836.152\,667$
electron–proton mass ratio	$m_\mathrm{e}/m_\mathrm{p}$	$5.446\,170\,10^{-4}$
proton magnetic moment	μ_p	$1.410\,607 \times 10^{-26}\,\mathrm{J\,T^{-1}}$ $2.792\,847\,\mu_\mathrm{N}$
proton g-factor	$g_\mathrm{p} = 2\mu_\mathrm{p}/\mu_\mathrm{N}$	$5.585\,695$
proton gyromagnetic ratio	$\gamma_\mathrm{p} = 2\mu_\mathrm{p}/\hbar$ $\gamma_\mathrm{p}/2\pi$	$2.675\,222 \times 10^{8}\,\mathrm{s^{-1}\,T^{-1}}$ $42.577\,481\,\mathrm{MHz\,T^{-1}}$
Rydberg constant	$R_\infty = \alpha^2 m_\mathrm{e} c/2h$	$10\,973\,731.569\,\mathrm{m^{-1}}$
Rydberg energy in electronvolts	$\mathrm{Ry} = R_\infty hc$	$2.179\,872 \times 10^{-18}\,\mathrm{J}$ $13.605\,692\,\mathrm{eV}$
Hartree energy in electronvolts	$E_\mathrm{h} = e^2/4\pi\epsilon_0 a_0$	$4.359\,744 \times 10^{-18}\,\mathrm{J}$ $27.211\,384\,\mathrm{eV}$

A.2 Relationships Among Units

The fundamental units of the SI system are meter (m), kilogram (kg), second (s), ampere (A), kelvin (K), mole (mol), and candela (Cd). In the next table derived units are expressed in terms of these fundamental ones – and other derived units.

$1\,\text{coulomb (C)} = 1\,\text{A s}$	$1\,\text{pascal (Pa)} = 1\,\text{N m}^{-2}$
$1\,\text{farad (F)} = 1\,\text{C V}^{-1} = 1\,\text{A}^2\,\text{s}^2\,\text{J}^{-1}$	$1\,\text{siemens (S)} = 1\,\Omega^{-1} = 1\,\text{A V}^{-1}$
$1\,\text{henry (H)} = 1\,\text{V s A}^{-1} = 1\,\text{J A}^{-2}$	$1\,\text{tesla (T)} = 1\,\text{Wb m}^{-2}$
$1\,\text{joule (J)} = 1\,\text{N m} = 1\,\text{kg m}^2\,\text{s}^{-2}$	$1\,\text{volt (V)} = 1\,\text{W A}^{-1}$
$1\,\text{newton (N)} = 1\,\text{kg m s}^{-2}$	$1\,\text{watt (W)} = 1\,\text{J s}^{-1}$
$1\,\text{ohm (}\Omega\text{)} = 1\,\text{V A}^{-1}$	$1\,\text{weber (Wb)} = 1\,\text{V s} = 1\,\text{J A}^{-1}$

It is immediately established that the unit of magnetic moment is $1\,\text{A m}^2 = 1\,\text{J T}^{-1}$.

The following SI prefixes are used to denote the multiples and fractions of the units:

Name	Symbol	Corresponding power of 10
yotta-	Y	10^{24}
zetta-	Z	10^{21}
exa-	E	10^{18}
peta-	P	10^{15}
tera-	T	10^{12}
giga-	G	10^{9}
mega-	M	10^{6}
kilo-	k	10^{3}
milli-	m	10^{-3}
micro-	μ	10^{-6}
nano-	n	10^{-9}
pico-	p	10^{-12}
femto-	f	10^{-15}
atto-	a	10^{-18}
zepto-	z	10^{-21}
yocto-	y	10^{-24}

Non-SI Units

Much of the solid-state physics literature continues to use CGS units. Some of the most frequently used units are:

$1\,\text{Å} = 10^{-10}\,\text{m} = 0.1\,\text{nm}\,,$	$1\,\text{erg} = 10^{-7}\,\text{J}\,,$
$1\,\text{atm} = 101\,325\,\text{Pa}\,,$	$1\,\text{bar} = 10^{5}\,\text{Pa}\,.$

The CGS unit of magnetic field strength is the oersted (Oe), while that of magnetic induction (flux density) and magnetization is the gauss (G). In SI, magnetic induction is given in teslas, while magnetic field strength and magnetization in A/m. Therefore

$$1\,\text{G} \mathrel{\widehat{=}} 10^{-4}\,\text{T}$$

when specifying the magnetic induction, however

$$1\,\mathrm{G} \mathrel{\widehat{=}} 10^3\,\mathrm{A/m}$$

when it comes to magnetization. The relationship between the units of magnetic flux is

$$1\,\mathrm{G\,cm^2} \mathrel{\widehat{=}} 10^{-8}\,\mathrm{Wb}\,.$$

The units of magnetic field strength are related by

$$1\,\mathrm{Oe} \mathrel{\widehat{=}} \frac{10^3}{4\pi}\,\mathrm{A/m} = 79.58\,\mathrm{A/m}\,.$$

When CGS units are used, equations for electromagnetic quantities are usually written in their nonrationalized form. The table below shows the conversion factors used in the transformation of the rationalized equations into nonrationalized ones that apply to the Gaussian quantities (denoted by *); $c^2 = 1/\epsilon_0\mu_0$.

Physical quantity	Conversion formula (CGS–SI)
electric charge	$q^* = (4\pi\epsilon_0)^{-1/2} q$
electric current	$j^* = (4\pi\epsilon_0)^{-1/2} j$
electric field	$\boldsymbol{E}^* = (4\pi\epsilon_0)^{1/2}\boldsymbol{E}$
electric displacement	$\boldsymbol{D}^* = (4\pi/\epsilon_0)^{1/2}\boldsymbol{D}$
magnetic field	$\boldsymbol{H}^* = (4\pi\mu_0)^{1/2}\boldsymbol{H}$
magnetic induction	$\boldsymbol{B}^* = (4\pi/\mu_0)^{1/2}\boldsymbol{B}$
scalar potential	$\varphi^* = (4\pi\epsilon_0)^{1/2}\varphi$
vector potential	$\boldsymbol{A}^* = (4\pi/\mu_0)^{1/2}\boldsymbol{A}$
resistivity	$\varrho^* = 4\pi\epsilon_0\varrho$
conductivity	$\sigma^* = (4\pi\epsilon_0)^{-1}\sigma$
permittivity	$\epsilon^* = \epsilon/\epsilon_0$
permeability	$\mu^* = \mu/\mu_0$
magnetic susceptibility	$\chi^* = \chi/4\pi$
magnetization	$\boldsymbol{M}^* = (\mu_0/4\pi)^{1/2}\boldsymbol{M}$
magnetic flux	$\Phi^* = (4\pi/\mu_0)^{1/2}\Phi$

For example, the expression $\omega_\mathrm{c} = eB/m_\mathrm{e}$ for cyclotron frequency goes over into

$$\omega_\mathrm{c} = \sqrt{\epsilon_0\mu_0}e^*B^*/m_\mathrm{e} = \frac{e^*B^*}{m_\mathrm{e}c}\,.$$

To convert electromagnetic SI units into CGS units, the next relationship (of mixed units) has to be used:

$$\sqrt{\frac{\mu_0}{4\pi}} = 0.1\,\frac{\sqrt{\mathrm{dyn}}}{\mathrm{A}}\,.$$

The nonrationalized form and CGS value of some electromagnetic constants are listed below.

Name of physical constant	Symbol	Value
elementary charge	e^*	$4.803\,24 \times 10^{-10}\,\mathrm{dyn}^{1/2}\,\mathrm{cm}$
Bohr radius	$a_0 = \hbar^2/m_\mathrm{e}e^{*2}$	$5.291\,772 \times 10^{-9}\,\mathrm{cm}$
Bohr magneton	$\mu_\mathrm{B}^* = e^*\hbar/2m_\mathrm{e}c$	$9.274\,009 \times 10^{-21}\,\mathrm{erg\,G}^{-1}$
fine-structure constant	$\alpha = e^{*2}/\hbar c$	$1/137.036$
flux quantum	$\Phi_0^* = hc/2e^*$	$2.067\,834 \times 10^{-7}\,\mathrm{G\,cm}^2$

Conversion Factors of Energy Equivalents

In solid-state physics energies are very often given in electronvolts: $1\,\mathrm{eV} = (e/\mathrm{C})\mathrm{J}$. Thermal energies are usually specified using the relation $E = k_\mathrm{B}T$, with the temperature given in kelvins, while magnetic field energies are commonly converted to field strengths given in teslas or gausses through $E = \mu_\mathrm{B}B$. In spectroscopy energy is often expressed in hertz units via $E = h\nu$ or in inverse (centi)meters, via $E = hc/\lambda$. The energies corresponding to these units are

$$
\begin{aligned}
E(1\,\mathrm{K}) &= (1\,\mathrm{K})k_\mathrm{B} = 1.380\,650 \times 10^{-23}\,\mathrm{J}\,,\\
E(1\,\mathrm{T}) &= (1\,\mathrm{T})\mu_\mathrm{B} = 9.274\,009 \times 10^{-24}\,\mathrm{J}\,,\\
E(1\,\mathrm{cm}^{-1}) &= (1\,\mathrm{cm}^{-1})hc = 1.986\,445 \times 10^{-23}\,\mathrm{J}\,,\\
E(1\,\mathrm{Hz}) &= (1\,\mathrm{Hz})h = 6.626\,069 \times 10^{-34}\,\mathrm{J}\,.
\end{aligned}
$$

The rydberg ($1\,\mathrm{Ry} = 1\,R_\infty hc$) and hartree ($1\,\mathrm{hartree} = 2\,\mathrm{Ry}$) units are widely used, too. The conversion factors relating energies given in the previous units are listed in the following table.

$$
\begin{aligned}
1\,\mathrm{eV} \mathrel{\widehat{=}}\ & 1.602\,176 \times 10^{-19}\,\mathrm{J}\\
& 1.160\,451 \times 10^{4}\,\mathrm{K}\\
& 1.727\,599 \times 10^{4}\,\mathrm{T}\\
& 7.349\,865 \times 10^{-2}\,\mathrm{Ry}\\
& 8.065\,544 \times 10^{3}\,\mathrm{cm}^{-1}\\
& 2.417\,989 \times 10^{14}\,\mathrm{Hz}
\end{aligned}
\qquad
\begin{aligned}
1\,\mathrm{K} \mathrel{\widehat{=}}\ & 1.380\,650 \times 10^{-23}\,\mathrm{J}\\
& 8.617\,343 \times 10^{-5}\,\mathrm{eV}\\
& 1.488\,731\,\mathrm{T}\\
& 6.333\,631 \times 10^{-6}\,\mathrm{Ry}\\
& 6.950\,356 \times 10^{-1}\,\mathrm{cm}^{-1}\\
& 2.083\,664 \times 10^{10}\,\mathrm{Hz}
\end{aligned}
$$

$$
\begin{aligned}
1\,\mathrm{T} \mathrel{\widehat{=}}\ & 9.274\,009 \times 10^{-24}\,\mathrm{J}\\
& 5.788\,382 \times 10^{-5}\,\mathrm{eV}\\
& 0.671\,713\,\mathrm{K}\\
& 4.254\,383 \times 10^{-6}\,\mathrm{Ry}\\
& 4.668\,645 \times 10^{-1}\,\mathrm{cm}^{-1}\\
& 1.399\,625 \times 10^{10}\,\mathrm{Hz}
\end{aligned}
\qquad
\begin{aligned}
1\,\mathrm{Ry} \mathrel{\widehat{=}}\ & 2.179\,872 \times 10^{-18}\,\mathrm{J}\\
& 13.605\,692\,\mathrm{eV}\\
& 1.578\,873 \times 10^{5}\,\mathrm{K}\\
& 2.350\,517 \times 10^{5}\,\mathrm{T}\\
& 1.097\,373 \times 10^{5}\,\mathrm{cm}^{-1}\\
& 3.289\,842 \times 10^{15}\,\mathrm{Hz}
\end{aligned}
$$

$1\,\mathrm{cm}^{-1} \mathrel{\widehat{=}}$	$1.986\,445 \times 10^{-23}\,\mathrm{J}$	$1\,\mathrm{Hz} \mathrel{\widehat{=}}$	$6.626\,069 \times 10^{-34}\,\mathrm{J}$
	$1.239\,842 \times 10^{-4}\,\mathrm{eV}$		$4.135\,667 \times 10^{-15}\,\mathrm{eV}$
	$1.438\,775\,\mathrm{K}$		$4.799\,237 \times 10^{-11}\,\mathrm{K}$
	$2.141\,949\,\mathrm{T}$		$7.144\,773 \times 10^{-11}\,\mathrm{T}$
	$9.112\,670 \times 10^{-6}\,\mathrm{Ry}$		$3.039\,660 \times 10^{-16}\,\mathrm{Ry}$
	$2.997\,925 \times 10^{10}\,\mathrm{Hz}$		$3.335\,641 \times 10^{-11}\,\mathrm{cm}^{-1}$

Binding and cohesive energies in solids are commonly given per atom or per mole. In addition to eV, Ry, and J, older tables also contain data given in calories. Its usual value is $1\,\mathrm{cal} = 4.1868\,\mathrm{J}$, however that of the thermochemical calory is $1\,\mathrm{cal} = 4.184\,\mathrm{J}$. Consequently

$$1\,\frac{\mathrm{eV}}{\mathrm{atom}} = 73.499\,\frac{\mathrm{mRy}}{\mathrm{atom}} = 96.4853\,\frac{\mathrm{kJ}}{\mathrm{mol}} = 23.05\,\frac{\mathrm{kcal}}{\mathrm{mol}}\,.$$

Reference

1. P. J. Mohr and B. N. Taylor, *2002 CODATA recommended values of the fundamental physical constants*, National Institute of Standards and Technology (2003). The data are available on the WWW at http://physics.nist.gov/cuu/Constants.

B

The Periodic Table of Elements

B.1 The Electron and Crystal Structures of Elements

The next table shows the electron structure for the elements of the periodic table as well as the Pearson symbol of their crystal structure that is stable at low temperatures and atmospheric pressure. When another crystal structure is stable at room temperature, its Pearson symbol is also listed.

Atomic number	Name of element	Chemical symbol	Electron structure	Crystal structure at 0 K	Crystal structure at 300 K
1	hydrogen	H	$1s^1$	cF4 (fcc)	gas
2	helium	He	$1s^2$	hP2 (hcp)	gas
3	lithium	Li	[He] $2s^1$	hP2 (hcp)	cI2 (bcc)
4	beryllium	Be	[He] $2s^2$	hP2 (hcp)	
5	boron	B	[He] $2s^2\ 2p^1$	hR105	
6	carbon	C	[He] $2s^2\ 2p^2$	hP4	
7	nitrogen	N	[He] $2s^2\ 2p^3$	cP8	gas
8	oxygen	O	[He] $2s^2\ 2p^4$	mC4	gas
9	fluorine	F	[He] $2s^2\ 2p^5$	mC6	gas
10	neon	Ne	[He] $2s^2\ 2p^6$	cF4 (fcc)	gas
11	sodium	Na	[Ne] $3s^1$	hP2 (hcp)	cI2 (bcc)
12	magnesium	Mg	[Ne] $3s^2$	hP2 (hcp)	
13	aluminum	Al	[Ne] $3s^2\ 3p^1$	cF4 (fcc)	
14	silicon	Si	[Ne] $3s^2\ 3p^2$	cF8	
15	phosphorus	P	[Ne] $3s^2\ 3p^3$	oC8	
16	sulfur	S	[Ne] $3s^2\ 3p^4$	oF128	
17	chlorine	Cl	[Ne] $3s^2\ 3p^5$	oC8	gas
18	argon	Ar	[Ne] $3s^2\ 3p^6$	cF4 (fcc)	gas
19	potassium	K	[Ar] $4s^1$	cI2 (bcc)	

Continued on the next page

Atomic number	Name of element	Chemical symbol	Electron structure	Crystal structure at 0 K	Crystal structure at 300 K
20	calcium	Ca	[Ar] $4s^2$	cF4 (fcc)	
21	scandium	Sc	[Ar] $4s^2\,3d^1$	hP2 (hcp)	
22	titanium	Ti	[Ar] $4s^2\,3d^2$	hP2 (hcp)	
23	vanadium	V	[Ar] $4s^2\,3d^3$	cI2 (bcc)	
24	chromium	Cr	[Ar] $4s^1\,3d^5$	cI2 (bcc)	
25	manganese	Mn	[Ar] $4s^2\,3d^5$	cI58	
26	iron	Fe	[Ar] $4s^2\,3d^6$	cI2 (bcc)	
27	cobalt	Co	[Ar] $4s^2\,3d^7$	hP2 (hcp)	
28	nickel	Ni	[Ar] $4s^2\,3d^8$	cF4 (fcc)	
29	copper	Cu	[Ar] $4s^1\,3d^{10}$	cF4 (fcc)	
30	zinc	Zn	[Ar] $4s^2\,3d^{10}$	hP2 (hcp)	
31	gallium	Ga	[Ar] $4s^2\,3d^{10}\,4p^1$	oC8	
32	germanium	Ge	[Ar] $4s^2\,3d^{10}\,4p^2$	cF8	
33	arsenic	As	[Ar] $4s^2\,3d^{10}\,4p^3$	hR2	
34	selenium	Se	[Ar] $4s^2\,3d^{10}\,4p^4$	hP3	
35	bromine	Br	[Ar] $4s^2\,3d^{10}\,4p^5$	oC8	liquid
36	krypton	Kr	[Ar] $4s^2\,3d^{10}\,4p^6$	cF4 (fcc)	gas
37	rubidium	Rb	[Kr] $5s^1$	cI2 (bcc)	
38	strontium	Sr	[Kr] $5s^2$	cF4 (fcc)	
39	yttrium	Y	[Kr] $5s^2\,4d^1$	hP2 (hcp)	
40	zirconium	Zr	[Kr] $5s^2\,4d^2$	hP2 (hcp)	
41	niobium	Nb	[Kr] $5s^1\,4d^4$	cI2 (bcc)	
42	molybdenum	Mo	[Kr] $5s^1\,4d^5$	cI2 (bcc)	
43	technetium	Tc	[Kr] $5s^2\,4d^5$	hP2 (hcp)	
44	ruthenium	Ru	[Kr] $5s^1\,4d^7$	hP2 (hcp)	
45	rhodium	Rh	[Kr] $5s^1\,4d^8$	cF4 (fcc)	
46	palladium	Pd	[Kr] $4d^{10}$	cF4 (fcc)	
47	silver	Ag	[Kr] $5s^1\,4d^{10}$	cF4 (fcc)	
48	cadmium	Cd	[Kr] $5s^2\,4d^{10}$	hP2 (hcp)	
49	indium	In	[Kr] $5s^2\,4d^{10}\,5p^1$	tI2	
50	tin	Sn	[Kr] $5s^2\,4d^{10}\,5p^2$	cF8	tI4
51	antimony	Sb	[Kr] $5s^2\,4d^{10}\,5p^3$	hR2	
52	tellurium	Te	[Kr] $5s^2\,4d^{10}\,5p^4$	hP3	
53	iodine	I	[Kr] $5s^2\,4d^{10}\,5p^5$	oC8	
54	xenon	Xe	[Kr] $5s^2\,4d^{10}\,5p^6$	cF4 (fcc)	gas
55	cesium	Cs	[Xe] $6s^1$	cI2 (bcc)	
56	barium	Ba	[Xe] $6s^2$	cI2 (bcc)	
57	lanthanum	La	[Xe] $6s^2\,5d^1$	hP4 (dhcp)	
58	cerium	Ce	[Xe] $6s^2\,4f^2$	cF4 (fcc)	hP4 (dhcp)
59	praseodymium	Pr	[Xe] $6s^2\,4f^3$	hP4 (dhcp)	
60	neodymium	Nd	[Xe] $6s^2\,4f^4$	hP4 (dhcp)	
61	promethium	Pm	[Xe] $6s^2\,4f^5$	hP4 (dhcp)	
62	samarium	Sm	[Xe] $6s^2\,4f^6$	hR3	

Continued on the next page

Atomic number	Name of element	Chemical symbol	Electron structure	Crystal structure at 0 K	Crystal structure at 300 K
63	europium	Eu	[Xe] $6s^2\,4f^7$	cI2 (bcc)	
64	gadolinium	Gd	[Xe] $6s^2\,4f^7\,5d^1$	hP2 (hcp)	
65	terbium	Tb	[Xe] $6s^2\,4f^9$	hP2 (hcp)	
66	dysprosium	Dy	[Xe] $6s^2\,4f^{10}$	oC4	hP2 (hcp)
67	holmium	Ho	[Xe] $6s^2\,4f^{11}$	hP2 (hcp)	
68	erbium	Er	[Xe] $6s^2\,4f^{12}$	hP2 (hcp)	
69	thulium	Tm	[Xe] $6s^2\,4f^{13}$	hP2 (hcp)	
70	ytterbium	Yb	[Xe] $6s^2\,4f^{14}$	hP2 (hcp)	cF4 (fcc)
71	lutetium	Lu	[Xe] $6s^2\,4f^{14}\,5d^1$	hP2 (hcp)	
72	hafnium	Hf	[Xe] $6s^2\,4f^{14}\,5d^2$	hP2 (hcp)	
73	tantalum	Ta	[Xe] $6s^2\,4f^{14}\,5d^3$	cI2 (bcc)	
74	tungsten	W	[Xe] $6s^2\,4f^{14}\,5d^4$	cI2 (bcc)	
75	rhenium	Re	[Xe] $6s^2\,4f^{14}\,5d^5$	hP2 (hcp)	
76	osmium	Os	[Xe] $6s^2\,4f^{14}\,5d^6$	hP2 (hcp)	
77	iridium	Ir	[Xe] $6s^2\,4f^{14}\,5d^7$	cF4 (fcc)	
78	platinum	Pt	[Xe] $6s^1\,4f^{14}\,5d^9$	cF4 (fcc)	
79	gold	Au	[Xe] $6s^1\,4f^{14}\,5d^{10}$	cF4 (fcc)	
80	mercury	Hg	[Xe] $6s^2\,4f^{14}\,5d^{10}$	hR1	liquid
81	thallium	Tl	[Xe] $6s^2\,4f^{14}\,5d^{10}\,6p^1$	hP2 (hcp)	
82	lead	Pb	[Xe] $6s^2\,4f^{14}\,5d^{10}\,6p^2$	cF4 (fcc)	
83	bismuth	Bi	[Xe] $6s^2\,4f^{14}\,5d^{10}\,6p^3$	hR2	
84	polonium	Po	[Xe] $6s^2\,4f^{14}\,5d^{10}\,6p^4$	cP1 (sc)	
85	astatine	At	[Xe] $6s^2\,4f^{14}\,5d^{10}\,6p^5$		
86	radon	Rn	[Xe] $6s^2\,4f^{14}\,5d^{10}\,6p^6$		gas
87	francium	Fr	[Rn] $7s^1$		
88	radium	Ra	[Rn] $7s^2$	cI2 (bcc)	
89	actinium	Ac	[Rn] $7s^2\,6d^1$	cF4 (fcc)	
90	thorium	Th	[Rn] $7s^2\,6d^2$	cF4 (fcc)	
91	protactinium	Pa	[Rn] $7s^2\,5f^2\,6d^1$	tI2	
92	uranium	U	[Rn] $7s^2\,5f^3\,6d^1$	oC4	
93	neptunium	Np	[Rn] $7s^2\,5f^4\,6d^1$	oP8	
94	plutonium	Pu	[Rn] $7s^2\,5f^6$	mP16	
95	americium	Am	[Rn] $7s^2\,5f^7$	hP4 (dhcp)	
96	curium	Cm	[Rn] $7s^2\,5f^7\,6d^1$	hP4 (dhcp)	
97	berkelium	Bk	[Rn] $7s^2\,5f^9$	hP4 (dhcp)	
98	californium	Cf	[Rn] $7s^2\,5f^{10}$	hP4 (dhcp)	
99	einsteinium	Es	[Rn] $7s^2\,5f^{11}$	hP4 (dhcp)	
100	fermium	Fm	[Rn] $7s^2\,5f^{12}$		
101	mendelevium	Md	[Rn] $7s^2\,5f^{13}$		
102	nobelium	No	[Rn] $7s^2\,5f^{14}$		
103	lawrencium	Lr	[Rn] $7s^2\,5f^{14}\,6d^1$		
104	rutherfordium	Rf	[Rn] $7s^2\,5f^{14}\,6d^2$		
105	dubnium	Db	[Rn] $7s^2\,5f^{14}\,6d^3$		

Continued on the next page

Atomic number	Name of element	Chemical symbol	Electron structure	Crystal structure at 0 K	Crystal structure at 300 K
106	seaborgium	Sg	[Rn] $7s^2\,5f^{14}\,6d^4$		
107	bohrium	Bh	[Rn] $7s^2\,5f^{14}\,6d^5$		
108	hassium	Hs	[Rn] $7s^2\,5f^{14}\,6d^6$		
109	meitnerium	Mt	[Rn] $7s^2\,5f^{14}\,6d^7$		
110	darmstadtium	Ds	[Rn] $7s^1\,5f^{14}\,6d^9$		
111	roentgenium	Rg	[Rn] $7s^1\,5f^{14}\,6d^{10}$		

B.2 Characteristic Temperatures of the Elements

The following table contains the melting point (in Celsius degrees), the Debye temperature (in kelvins), and the critical temperature of the superconducting or magnetic phase transition (in kelvins) for the elements of the periodic table. The symbols S, F, and AF denote superconducting, ferromagnetic, and antiferromagnetic phases, respectively. When more than one magnetic phases are possible, only the transition temperature to the ground-state structure is given. The symbol F does not mean that moments are rigorously parallel, only that the material possesses a finite net magnetization.

Atomic number	Name of element	Chemical symbol	Melting point (°C)	Θ_D (K)	Ordered phase	Transition temperature
1	hydrogen	H	−259	105		
2	helium	He		26		
3	lithium	Li	180	344		
4	beryllium	Be	1287	1440	S	$T_c = 0.026\,K$
5	boron	B	2075	1315		
6	carbon	C	3825	420		
7	nitrogen	N	−210	68		
8	oxygen	O	−218	91		
9	fluorine	F	−220			
10	neon	Ne	−249	75		
11	sodium	Na	98	158		
12	magnesium	Mg	650	400		
13	aluminum	Al	660	428	S	$T_c = 1.175\,K$
14	silicon	Si	1414	640		
15	phosphorus	P	44	193		
16	sulfur	S	115	250		
17	chlorine	Cl	−101	115		
18	argon	Ar	−189	93		
19	potassium	K	63	91		

Continued on the next page

Atomic number	Name of element	Chemical symbol	Melting point (°C)	Θ_D (K)	Ordered phase	Transition temperature
20	calcium	Ca	842	230		
21	scandium	Sc	1541	360		
22	titanium	Ti	1668	420	S	$T_c = 0.40\,K$
23	vanadium	V	1910	380	S	$T_c = 5.46\,K$
24	chromium	Cr	1907	630	AF	$T_N = 311\,K$
25	manganese	Mn	1246	410	AF	$T_N = 100\,K$
26	iron	Fe	1538	467	F	$T_C = 1043\,K$
27	cobalt	Co	1495	445	F	$T_C = 1388\,K$
28	nickel	Ni	1455	450	F	$T_C = 627\,K$
29	copper	Cu	1085	343		
30	zinc	Zn	420	327	S	$T_c = 0.86\,K$
31	gallium	Ga	30	320	S	$T_c = 1.08\,K$
32	germanium	Ge	938	370		
33	arsenic	As	817	282		
34	selenium	Se	221	90		
35	bromine	Br	−7			
36	krypton	Kr	−157	72		
37	rubidium	Rb	39	56		
38	strontium	Sr	777	147		
39	yttrium	Y	1522	280		
40	zirconium	Zr	1855	291	S	$T_c = 0.63\,K$
41	niobium	Nb	2477	275	S	$T_c = 9.25\,K$
42	molybdenum	Mo	2623	450	S	$T_c = 0.92\,K$
43	technetium	Tc	2157	351	S	$T_c = 7.8\,K$
44	ruthenium	Ru	2334	600	S	$T_c = 0.49\,K$
45	rhodium	Rh	1964	480	S	$T_c = 0.035\,mK$
46	palladium	Pd	1555	274		
47	silver	Ag	962	225		
48	cadmium	Cd	321	209	S	$T_c = 0.52\,K$
49	indium	In	157	108	S	$T_c = 3.41\,K$
50	tin	Sn	232	199	S	$T_c = 3.72\,K$
51	antimony	Sb	631	211		
52	tellurium	Te	450	153		
53	iodine	I	114	106		
54	xenon	Xe	−112	64		
55	cesium	Cs	28	38		
56	barium	Ba	727	110		
57	lanthanum	La	920	142	S	$T_c = 5\,K$
58	cerium	Ce	798	146	AF	$T_N = 12.5\,K$

Continued on the next page

Atomic number	Name of element	Chemical symbol	Melting point (°C)	Θ_D (K)	Ordered phase	Transition temperature
59	praseodymium	Pr	931	85	AF	$T_N = 0.03\,K$
60	neodymium	Nd	1016	159	AF	$T_N = 6\,K$
61	promethium	Pm	1042	158		
62	samarium	Sm	1074	116	AF	$T_N = 14.0\,K$
63	europium	Eu	822	127	AF	$T_N = 90.4\,K$
64	gadolinium	Gd	1313	195	F	$T_C = 293\,K$
65	terbium	Tb	1356	150	F	$T_C = 220\,K$
66	dysprosium	Dy	1412	210	F	$T_C = 90\,K$
67	holmium	Ho	1474	114	F	$T_C = 20\,K$
68	erbium	Er	1529	134	F	$T_C = 18\,K$
69	thulium	Tm	1545	127	F	$T_C = 32\,K$
70	ytterbium	Yb	824	118		
71	lutetium	Lu	1663	210	S	$T_c = 0.1\,K$
72	hafnium	Hf	2233	252	S	$T_c = 0.13\,K$
73	tantalum	Ta	3017	240	S	$T_c = 4.47\,K$
74	tungsten	W	3422	400	S	$T_c = 0.02\,K$
75	rhenium	Re	3186	430	S	$T_c = 1.70\,K$
76	osmium	Os	3033	500	S	$T_c = 0.66\,K$
77	iridium	Ir	2446	420	S	$T_c = 0.11\,K$
78	platinum	Pt	1768	240		
79	gold	Au	1064	165		
80	mercury	Hg	−39	72	S	$T_c = 4.15\,K$
81	thallium	Tl	304	78	S	$T_c = 2.38\,K$
82	lead	Pb	327	105	S	$T_c = 7.20\,K$
83	bismuth	Bi	271	119		
84	polonium	Po	254	81		
85	astatine	At	302			
86	radon	Rn	−71			
87	francium	Fr	27			
88	radium	Ra	696	89		
89	actinium	Ac	1051	124		
90	thorium	Th	1750	170	S	$T_c = 1.37\,K$
91	protactinium	Pa	1572	159	S	$T_c = 1.4\,K$
92	uranium	U	1135	207	S	$T_c = 0.68\,K$
93	neptunium	Np	644	121		
94	plutonium	Pu	640	171		
95	americium	Am	1176		S	$T_c = 0.60\,K$
96	curium	Cm	1345			

References

1. *CRC Handbook of Chemistry and Physics*, Editor-in-Chief D. R. Lide, 85th Edition, CRC Press, Boca Raton (2004).
2. http://www.webelements.com
3. *Springer Handbook of Condensed Matter and Materials Data*, Editors: W. Martienssen and H. Warlimont, Springer-Verlag, Berlin (2005).

C

Mathematical Formulas

C.1 Fourier Transforms

When surface effects are neglected and only bulk properties are examined in macroscopic crystalline samples, periodic boundary conditions are frequently applied, since the Fourier components that appear in the Fourier series or Fourier integral representation of the position-dependent quantities are often easier to determine. Nevertheless, owing to the invariance of crystals under discrete translations, functions that show the same periodicity as the crystal lattice and functions defined at the vertices of the crystal lattice are frequently encountered, too. It is often more convenient to specify them using the Fourier components associated with the vectors defined in the reciprocal lattice. The most important formulas of such functions are listed in the present section.

C.1.1 Fourier Transform of Continuous Functions

A periodic function of period L ($f(x+L) = f(x)$) can be expanded into a Fourier series as

$$f(x) = \tfrac{1}{2}a_0 + \sum_{n=0}^{\infty}\left[a_n \cos\left(\frac{2\pi n}{L}x\right) + b_n \sin\left(\frac{2\pi n}{L}x\right)\right]. \tag{C.1.1}$$

Making use of the orthogonality relation of trigonometric functions, it can be shown that the coefficients are given by the integrals

$$a_n = \frac{2}{L}\int_{-L/2}^{L/2} f(x)\cos\left(\frac{2\pi n}{L}x\right)\mathrm{d}x, \qquad b_n = \frac{2}{L}\int_{-L/2}^{L/2} f(x)\sin\left(\frac{2\pi n}{L}x\right)\mathrm{d}x\,. \tag{C.1.2}$$

It is often more convenient to use exponential functions:

$$f(x) = \sum_{n=-\infty}^{\infty} f_n \mathrm{e}^{2\pi \mathrm{i} n x/L}, \tag{C.1.3}$$

where

$$f_n = \frac{1}{L} \int_{-L/2}^{L/2} f(x) \mathrm{e}^{-2\pi \mathrm{i} n x/L} \, \mathrm{d}x \,. \tag{C.1.4}$$

This can be demonstrated directly by exploiting the completeness relation

$$\begin{aligned} \sum_{n=-\infty}^{\infty} \mathrm{e}^{2\pi \mathrm{i} n x/L} &= \lim_{N\to\infty} \sum_{n=-N}^{N} \mathrm{e}^{2\pi \mathrm{i} n x/L} \\ &= \lim_{N\to\infty} \frac{\sin[\pi(2N+1)x/L]}{\sin(\pi x/L)} = L\delta(x) \end{aligned} \tag{C.1.5}$$

and the orthogonality relation

$$\int_{-L/2}^{L/2} \mathrm{e}^{-2\pi \mathrm{i}(n-n')x/L} \, \mathrm{d}x = \frac{\sin[\pi(n-n')]}{\pi(n-n')/L} = L\delta_{n,n'} \,. \tag{C.1.6}$$

The coefficients f_n of the Fourier series make up the Fourier spectrum of function f.

The Fourier series representation is straightforward to generalize to functions defined in d-dimensional space, provided they satisfy periodic boundary conditions on hypercubes (or even more generally, hyperparallelepipeds) of volume L^d – and are repeated periodically outside it. For simplicity, consider a function f defined inside a three-dimensional general parallelepiped of edges $N_1\boldsymbol{a}_1$, $N_2\boldsymbol{a}_2$, $N_3\boldsymbol{a}_3$, and volume V that satisfies the periodic boundary conditions

$$f(\boldsymbol{r} + N_i\boldsymbol{a}_i) = f(\boldsymbol{r}) \,, \qquad i = 1,2,3 \tag{C.1.7}$$

on and beyond the boundaries. The Fourier series can then be written as

$$f(\boldsymbol{r}) = \frac{1}{V} \sum_{\boldsymbol{k}} \hat{f}(\boldsymbol{k}) \, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} \,. \tag{C.1.8}$$

On account of the periodic boundary conditions, the allowed vectors $\boldsymbol{k}$ are most easily expressed in terms of the primitive vectors $\boldsymbol{b}_i$ of the reciprocal lattice:

$$\boldsymbol{k} = \sum_{i=1}^{3} \frac{m_i}{N_i} \boldsymbol{b}_i \,, \tag{C.1.9}$$

where the m_i are arbitrary integers. Recall that the primitive vectors of the direct and reciprocal lattices are related by (5.2.13).

The explicit form of the Fourier coefficient $\hat{f}(\boldsymbol{k})$ can be derived either using the generalization

$$\sum_{\boldsymbol{k}} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{r}-\boldsymbol{r}')} = V\delta(\boldsymbol{r}-\boldsymbol{r}') \tag{C.1.10}$$

of the completeness relation (C.1.5), or the orthogonality relation

$$\int_V \mathrm{e}^{-\mathrm{i}(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{r}}\mathrm{d}\boldsymbol{r} = V\delta_{\boldsymbol{k},\boldsymbol{k}'}\,, \tag{C.1.11}$$

leading to

$$\hat{f}(\boldsymbol{k}) = \int_V f(\boldsymbol{r})\,\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\,\mathrm{d}\boldsymbol{r}\,. \tag{C.1.12}$$

It is easily seen that the convention used in the one-dimensional case is recovered if instead of (C.1.8) the Fourier series is defined as

$$f(\boldsymbol{r}) = \sum_{\boldsymbol{k}} \hat{f}(\boldsymbol{k})\,\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\,, \tag{C.1.13}$$

and consequently the Fourier coefficients are given by

$$\hat{f}(\boldsymbol{k}) = \frac{1}{V}\int_V f(\boldsymbol{r})\,\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\,\mathrm{d}\boldsymbol{r} \tag{C.1.14}$$

instead of (C.1.12). The rationale behind choosing a different convention is that in sufficiently large samples, where discrete sums are replaced by continuous integrals, the obtained formulas are independent of the sample volume in the $V \to \infty$ limit. Since each vector $\boldsymbol{k}$ in the primitive cell of the reciprocal lattice is associated with a volume $(2\pi)^3/V$, the sum over the $\boldsymbol{k}$ vectors can be replaced by an integral, using the formal substitution

$$\sum_{\boldsymbol{k}} \to \frac{V}{(2\pi)^3}\int \mathrm{d}\boldsymbol{k}\,. \tag{C.1.15}$$

Then the Fourier integral representation of an arbitrary function $f(\boldsymbol{r})$ defined on the whole space is, by definition,

$$f(\boldsymbol{r}) = \frac{1}{(2\pi)^3}\int \hat{f}(\boldsymbol{k})\,\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\,\mathrm{d}\boldsymbol{k}\,, \tag{C.1.16}$$

where

$$\hat{f}(\boldsymbol{k}) = \int f(\boldsymbol{r})\,\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\,\mathrm{d}\boldsymbol{r}\,, \tag{C.1.17}$$

as in the $V \to \infty$ limit

$$\int \frac{\mathrm{d}\boldsymbol{k}}{(2\pi)^3}\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{r}-\boldsymbol{r}')} = \delta(\boldsymbol{r}-\boldsymbol{r}')\,, \tag{C.1.18}$$

and

$$\int \mathrm{e}^{-\mathrm{i}(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{r}}\mathrm{d}\boldsymbol{r} = (2\pi)^3\delta(\boldsymbol{k}-\boldsymbol{k}')\,. \tag{C.1.19}$$

The function $\hat{f}(\boldsymbol{k})$ is the Fourier transform of $f(\boldsymbol{r})$ and (C.1.16) defines the inverse Fourier transform.

More generally, the Fourier transform of a function defined in d-dimensional space is given in the space of the d-dimensional $\boldsymbol{k}$ vectors as

$$\hat{f}(\boldsymbol{k}) = \int f(\boldsymbol{r})\, \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}\, \mathrm{d}^d x\,, \tag{C.1.20}$$

and the inverse Fourier transform is defined by

$$f(\boldsymbol{x}) = \frac{1}{(2\pi)^d} \int \hat{f}(\boldsymbol{k})\, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}\, \mathrm{d}^d k\,. \tag{C.1.21}$$

The completeness and orthogonality relations then take the form

$$\int \frac{\mathrm{d}^d k}{(2\pi)^d} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{x}-\boldsymbol{x}')} = \delta^{(d)}(\boldsymbol{x}-\boldsymbol{x}') \tag{C.1.22}$$

and

$$\int \mathrm{e}^{-\mathrm{i}(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{x}} \mathrm{d}^d x = (2\pi)^d \delta^{(d)}(\boldsymbol{k}-\boldsymbol{k}')\,. \tag{C.1.23}$$

In quantum mechanics, a common choice for the Fourier transform of the function $f(\boldsymbol{r})$ is

$$\hat{f}(\boldsymbol{k}) = \frac{1}{(2\pi)^{d/2}} \int_{-\infty}^{\infty} f(\boldsymbol{x}) \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}\, \mathrm{d}^d x\,, \tag{C.1.24}$$

and the inverse Fourier transform is then

$$f(\boldsymbol{x}) = \frac{1}{(2\pi)^{d/2}} \int_{-\infty}^{\infty} \hat{f}(\boldsymbol{k}) \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}\, \mathrm{d}^d k\,. \tag{C.1.25}$$

With this choice

$$\int \left|f(\boldsymbol{x})\right|^2 \mathrm{d}^d x = \int \left|\hat{f}(\boldsymbol{k})\right|^2 \mathrm{d}^d k\,, \tag{C.1.26}$$

or more generally

$$\int f^*(\boldsymbol{x}) g(\boldsymbol{x}) \mathrm{d}^d x = \int \hat{f}^*(\boldsymbol{k}) \hat{g}(\boldsymbol{k}) \mathrm{d}^d k\,, \tag{C.1.27}$$

which indicates that the Fourier transform is a unitary transformation in the space of square integrable functions that preserves lengths and inner products.

Another convention is used for the time variable. The Fourier transform of an arbitrary time-dependent function $f(t)$ is defined as

$$\hat{f}(\omega) = \int_{-\infty}^{\infty} f(t)\, \mathrm{e}^{\mathrm{i}\omega t}\, \mathrm{d}t\,, \tag{C.1.28}$$

and the inverse transform as

$$f(t) = \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} \hat{f}(\omega)\, \mathrm{e}^{-\mathrm{i}\omega t}\, \mathrm{d}\omega\,. \tag{C.1.29}$$

Therefore the following formula is used for space- and time-dependent functions that satisfy periodic boundary conditions at the boundaries of a sample of volume V:

$$\hat{f}(\boldsymbol{k}, \omega) = \int\limits_{V} \mathrm{d}\boldsymbol{r} \int\limits_{-\infty}^{\infty} \mathrm{d}t\, f(\boldsymbol{r}, t)\, \mathrm{e}^{-\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{r} - \omega t)}\,, \tag{C.1.30}$$

and

$$f(\boldsymbol{r}, t) = \frac{1}{V} \sum_{\boldsymbol{k}} \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} \mathrm{d}\omega\, \hat{f}(\boldsymbol{k}, \omega)\, \mathrm{e}^{\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{r} - \omega t)}\,, \tag{C.1.31}$$

while for samples of infinite extent

$$\hat{f}(\boldsymbol{k}, \omega) = \int \mathrm{d}\boldsymbol{r} \int\limits_{-\infty}^{\infty} \mathrm{d}t\, f(\boldsymbol{r}, t)\, \mathrm{e}^{-\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{r} - \omega t)}\,, \tag{C.1.32}$$

and

$$f(\boldsymbol{r}, t) = \frac{1}{(2\pi)^4} \int \mathrm{d}\boldsymbol{k} \int\limits_{-\infty}^{\infty} \mathrm{d}\omega\, \hat{f}(\boldsymbol{k}, \omega)\, \mathrm{e}^{\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{r} - \omega t)}\,. \tag{C.1.33}$$

These formulas can be applied to lattice-periodic functions, whose values inside the primitive cell of volume v spanned by the vectors $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$ are repeated with the periodicity of the lattice – in other words, for each translation vector $\boldsymbol{t}_n$ that can be written in the form (5.1.1),

$$f(\boldsymbol{r} + \boldsymbol{t}_n) = f(\boldsymbol{r})\,. \tag{C.1.34}$$

Since condition (C.1.7) is now met by the choice $N_1 = N_2 = N_3 = 1$, the vectors $\boldsymbol{k}$ appearing in the Fourier representation are the same as the vectors $\boldsymbol{G}$ of the reciprocal lattice, hence the Fourier transform of $f(\boldsymbol{r})$ is

$$\hat{f}(\boldsymbol{G}) = \int\limits_{v} f(\boldsymbol{r})\, \mathrm{e}^{-\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{r}}\, \mathrm{d}\boldsymbol{r}\,, \tag{C.1.35}$$

while the inverse transform is

$$f(\boldsymbol{r}) = \frac{1}{v} \sum_{\boldsymbol{G}} \hat{f}(\boldsymbol{G})\, \mathrm{e}^{\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{r}}\,. \tag{C.1.36}$$

C.1.2 Fourier Transform of Functions Defined at Lattice Points

Functions defined at the vertices of a discrete lattice are frequently used in solid-state physics. Consider a discrete lattice of volume V, with N lattice points. When the function $f(\boldsymbol{R}_i)$ is subject to periodic boundary conditions, it can be represented as

$$f(\boldsymbol{R}_i) = \frac{1}{N} \sum_{\boldsymbol{k}} \hat{f}(\boldsymbol{k})\, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}_i}\,, \tag{C.1.37}$$

and the Fourier transform $\hat{f}(\boldsymbol{k})$ can be written as

$$\hat{f}(\boldsymbol{k}) = \sum_{\boldsymbol{R}_i} f(\boldsymbol{R}_i)\, \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}_i}\,. \tag{C.1.38}$$

Since $\hat{f}(\boldsymbol{k}) = \hat{f}(\boldsymbol{k}+\boldsymbol{G})$ for any vector $\boldsymbol{G}$ of the reciprocal lattice, it is sufficient to consider one vector $\boldsymbol{k}$ in each set of equivalent vectors (which differ by a reciprocal-lattice vector) – in other words, the vectors $\boldsymbol{k}$ given in (C.1.9) are defined within the primitive cell or Brillouin zone of the reciprocal lattice. The number of allowed vectors $\boldsymbol{k}$ is just N.

To justify the previous formulas, we shall demonstrate that in the limit where the number of lattice points is large,

$$\frac{1}{N} \sum_{\boldsymbol{R}_i} \mathrm{e}^{-\mathrm{i}(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{R}_i} = \sum_{\boldsymbol{G}} \delta_{\boldsymbol{k}-\boldsymbol{k}',\boldsymbol{G}}\,. \tag{C.1.39}$$

The equality obviously holds when $\boldsymbol{k}-\boldsymbol{k}'$ is the same as a vector $\boldsymbol{G}$ of the reciprocal lattice, since each of the N terms in the sum

$$\sum_{\boldsymbol{R}_i} \mathrm{e}^{-\mathrm{i}(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{R}_i} \tag{C.1.40}$$

is then unity. Otherwise the phase factors cancel out to a good approximation. This cancellation can be most easily demonstrated in the case where the crystal contains odd numbers of lattice points along the direction of each primitive vector ($\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$). Expressed in terms of the primitive vectors $\boldsymbol{b}_1$, $\boldsymbol{b}_2$, $\boldsymbol{b}_3$ of the reciprocal lattice, $\boldsymbol{k}-\boldsymbol{k}'$ is

$$\boldsymbol{k}-\boldsymbol{k}' = \sum_i (k_i - k_i')\boldsymbol{b}_i\,, \tag{C.1.41}$$

and so the sum in question reads

$$\frac{1}{N} \sum_{n_1=-N_1}^{N_1} \sum_{n_2=-N_2}^{N_2} \sum_{n_3=-N_3}^{N_3} \exp\{-2\pi\mathrm{i}[(k_1-k_1')n_1 + (k_2-k_2')n_2 + (k_3-k_3')n_3]\}\,. \tag{C.1.42}$$

Performing the sum separately along the three directions,

$$\begin{aligned}
&\frac{1}{2N_i+1}\sum_{n_i=-N_i}^{N_i}\exp[-2\pi\mathrm{i}(k_i-k_i')n_i] \\
&\quad=\frac{\exp[2\pi\mathrm{i}(k_i-k_i')N_i]}{2N_i+1}\sum_{n_i=0}^{2N_i}\exp[-2\pi\mathrm{i}(k_i-k_i')n_i] \\
&\quad=\frac{\exp[2\pi\mathrm{i}(k_i-k_i')N_i]}{2N_i+1}\left[\frac{\exp[-2\pi\mathrm{i}(k_i-k_i')(2N_i+1)]-1}{\exp[-2\pi\mathrm{i}(k_i-k_i')]-1}\right] \\
&\quad=\frac{\sin[\pi(k_i-k_i')(2N_i+1)]}{(2N_i+1)\sin[\pi(k_i-k_i')]}\,. \qquad \text{(C.1.43)}
\end{aligned}$$

Wherever the denominator is zero, i.e., when $k_i - k_i' = 0, \pm 1, \pm 2, \ldots$ (in other words, $\boldsymbol{k} - \boldsymbol{k}'$ is a vector of the reciprocal lattice), the expression tends to 1 in the limit $N \to \infty$. Everywhere else it vanishes in the same limit.

Since the allowed vectors $\boldsymbol{k}$ are distributed continuously in the $N \to \infty$ limit, the relation (C.1.39) can be recast in the equivalent form

$$\frac{1}{N}\sum_{\boldsymbol{R}_i}\mathrm{e}^{-\mathrm{i}(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{R}_i}=\frac{(2\pi)^3}{V}\sum_{\boldsymbol{G}}\delta(\boldsymbol{k}-\boldsymbol{k}'-\boldsymbol{G})\,. \qquad \text{(C.1.44)}$$

In one dimension this can be rewritten as

$$\frac{1}{N}\sum_{R_n}\mathrm{e}^{-\mathrm{i}(k-k')\cdot R_n}=\frac{2\pi}{L}\sum_{h}\delta(k-k'-G_h)\,, \qquad \text{(C.1.45)}$$

where $R_n = na$ and $G_h = (2\pi/a)h$. Owing to the properties of the delta function, this is equivalent to

$$\sum_{n=-\infty}^{\infty}\mathrm{e}^{-2\pi\mathrm{i}nx}=\sum_{h=-\infty}^{\infty}\delta(x-h)\,. \qquad \text{(C.1.46)}$$

It can be shown along the same lines that for the sum over the discrete vectors $\boldsymbol{k}$ in the Brillouin zone

$$\frac{1}{N}\sum_{\boldsymbol{k}\in\mathrm{BZ}}\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{R}_i-\boldsymbol{R}_j)}=\delta_{\boldsymbol{R}_i,\boldsymbol{R}_j}\,. \qquad \text{(C.1.47)}$$

In the $N \to \infty$ limit, where the sum over the reciprocal lattice can be replaced by an integral,

$$\frac{V}{N}\frac{1}{(2\pi)^3}\int_{\boldsymbol{k}\in\mathrm{BZ}}\mathrm{d}\boldsymbol{k}\,\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{R}_i-\boldsymbol{R}_j)}=\delta_{\boldsymbol{R}_i,\boldsymbol{R}_j}\,. \qquad \text{(C.1.48)}$$

Naturally, integration is once again over the primitive cell or Brillouin zone of the reciprocal lattice.

A different normalization of Fourier transform can be chosen for functions defined on a lattice, too. The alternative

$$\begin{aligned} f(\boldsymbol{R}_i) &= \frac{1}{\sqrt{N}} \sum_{\boldsymbol{k}} \hat{f}(\boldsymbol{k})\, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}_i}\,, \\ \hat{f}(\boldsymbol{k}) &= \frac{1}{\sqrt{N}} \sum_{\boldsymbol{R}_i} f(\boldsymbol{R}_i)\, \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}_i}\,, \end{aligned} \tag{C.1.49}$$

which is commonly used in the literature, is also extensively applied in the present book.

C.1.3 Fourier Transform of Some Simple Functions

Owing to the periodicity of $\sin\varphi$, the function $\exp(\mathrm{i}z\sin\varphi)$ is also periodic with period 2π. Its Fourier decomposition is

$$\mathrm{e}^{\mathrm{i}z\sin\varphi} = \sum_{n=-\infty}^{\infty} J_n(z)\, \mathrm{e}^{\mathrm{i}n\varphi}\,, \tag{C.1.50}$$

where $J_n(z)$ is the Bessel function of order n. Similarly,

$$\mathrm{e}^{\mathrm{i}z\cos\varphi} = \sum_{n=-\infty}^{\infty} \mathrm{i}^n J_n(z)\, \mathrm{e}^{\mathrm{i}n\varphi}\,. \tag{C.1.51}$$

The sawtooth wave is defined to take the value $2x/L$ over the interval $-L/2 < x < L/2$, which is then repeated with period L in both directions. Its Fourier representation is

$$f(x) = -\frac{2}{\pi} \sum_{n=1}^{\infty} \frac{(-1)^n}{n} \sin\frac{2\pi n x}{L}\,. \tag{C.1.52}$$

The Fourier transform of the Heaviside step function

$$\theta(x) = \begin{cases} 1 & x > 0\,, \\ \frac{1}{2} & x = 0\,, \\ 0 & x < 0 \end{cases} \tag{C.1.53}$$

is

$$\hat{\theta}(k) = \pi\delta(k) - \frac{\mathrm{i}}{k}\,. \tag{C.1.54}$$

To determine the Fourier transform of the top-hat function

$$\mathit{\Pi}(x) = \begin{cases} 1/a & -a/2 \le x \le a/2\,, \\ 0 & |x| > a/2 \end{cases} \tag{C.1.55}$$

one has to exploit the relation

$$\Pi(x) = \frac{1}{a}[\theta(x + a/2) - \theta(x - a/2)] \tag{C.1.56}$$

and the property that the Fourier transform of $\theta(x - x_0)$ differs from that of $\theta(x)$ by a factor $\exp(-\mathrm{i}kx_0)$. The result is

$$\hat{\Pi}(k) = \frac{\sin(ka/2)}{ka/2}\,. \tag{C.1.57}$$

The Fourier transform of the Gaussian function

$$f(x) = \mathrm{e}^{-ax^2} \tag{C.1.58}$$

is another Gaussian function,

$$\hat{f}(k) = \sqrt{\pi/a}\,\mathrm{e}^{-k^2/4a}\,. \tag{C.1.59}$$

The Fourier transform of the Lorentzian function

$$f(x) = \frac{1}{\pi}\frac{\Gamma}{(x - x_0)^2 + \Gamma^2} \tag{C.1.60}$$

contains an exponentially decaying term,

$$\hat{f}(k) = \mathrm{e}^{-\mathrm{i}kx_0 - \Gamma|k|}\,. \tag{C.1.61}$$

Because of its slow decay, the Coulomb potential $1/r$ does not have an unambiguous Fourier transform, but the Yukawa potential, which contains an additional exponentially decaying factor, does:

$$\int \frac{\mathrm{e}^{-\alpha r}}{r}\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\,\mathrm{d}\boldsymbol{r} = \frac{4\pi}{k^2 + \alpha^2}\,. \tag{C.1.62}$$

The relationship

$$\int \frac{1}{r}\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\,\mathrm{d}\boldsymbol{r} = \frac{4\pi}{k^2} \tag{C.1.63}$$

can be interpreted as a limit of the preceding formula. The inverse Fourier transform leads to

$$\frac{1}{r} = \frac{4\pi}{V}\sum_{\boldsymbol{k}}\frac{1}{k^2}\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\,. \tag{C.1.64}$$

Changing to a continuous variable,

$$\frac{1}{r} = \frac{1}{2\pi^2}\int \frac{1}{k^2}\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\,\mathrm{d}\boldsymbol{k}\,. \tag{C.1.65}$$

The following inverse Fourier transform is related to the Green function of particles with a k^2 dispersion relation:

$$\frac{1}{V}\sum_{\boldsymbol{k}}\frac{\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}}{k^2 - a^2} = \frac{\cos ar}{4\pi r}\,. \tag{C.1.66}$$

C.2 Some Useful Integrals

Below we shall present some definite integrals of particular importance in solid-state physics. The special functions in terms of which certain integrals are expressed will be introduced and evaluated in the next section.

C.2.1 Integrals Containing Exponential Functions

Integrals that contain the product of an exponential and a power-law function can very generally be rewritten in the form

$$\int_0^\infty \mathrm{e}^{-\mu x} x^{\nu-1}\,\mathrm{d}x = \frac{1}{\mu^\nu}\Gamma(\nu) \qquad \mathrm{Re}\,\mu > 0\,,\ \ \mathrm{Re}\,\nu > 0\,. \tag{C.2.1}$$

In the special case when the power is an integer or a half-integer,

$$\int_0^\infty \mathrm{e}^{-x} x^{n-1}\,\mathrm{d}x = \Gamma(n)\,, \tag{C.2.2}$$

$$\int_0^\infty \mathrm{e}^{-\alpha x}\sqrt{x}\,\mathrm{d}x = \frac{\sqrt{\pi}}{2\alpha^{3/2}}\,. \tag{C.2.3}$$

When x^2 appears in the exponent,

$$\int_0^\infty \mathrm{e}^{-\alpha x^2} x^{2n+1}\,\mathrm{d}x = \frac{n!}{2\alpha^{n+1}}\,, \tag{C.2.4}$$

$$\int_0^\infty \mathrm{e}^{-\alpha x^2} x^{2n}\,\mathrm{d}x = \frac{(2n-1)!!}{2^{n+1}}\sqrt{\frac{\pi}{\alpha^{2n+1}}} \tag{C.2.5}$$

for odd and even powers of x, while for $n = 0$

$$\int_0^\infty \mathrm{e}^{-\alpha x^2}\,\mathrm{d}x = \frac{1}{2}\sqrt{\frac{\pi}{\alpha}}\,. \tag{C.2.6}$$

When the argument of the exponent is imaginary,

$$\int_0^\infty \mathrm{e}^{-\mathrm{i}\alpha x^2}\,\mathrm{d}x = \mathrm{e}^{-\mathrm{i}\pi/4}\sqrt{\frac{\pi}{\alpha}}\,. \tag{C.2.7}$$

C.2.2 Integrals Containing the Bose Function

When the chemical potential is zero, the definite integral over the interval $(0,\infty)$ of the product of the Bose function $1/(\mathrm{e}^x-1)$ and an arbitrary positive power of x can be determined exactly:

$$\int_0^\infty \frac{x^{\nu-1}}{\mathrm{e}^x-1}\,\mathrm{d}x = \sum_{k=1}^\infty \int_0^\infty x^{\nu-1}\,\mathrm{e}^{-kx}\,\mathrm{d}x = \sum_{k=1}^\infty \frac{1}{k^\nu}\Gamma(\nu) = \Gamma(\nu)\zeta(\nu) \tag{C.2.8}$$

if $\nu > 1$. When ν is an integer,

$$\int_0^\infty \frac{x^n}{\mathrm{e}^x-1}\,\mathrm{d}x = \sum_{k=1}^\infty \int_0^\infty x^n\,\mathrm{e}^{-kx}\,\mathrm{d}x = n!\sum_{k=1}^\infty \frac{1}{k^{n+1}} = n!\zeta(n+1)\,. \tag{C.2.9}$$

For odd powers, the integral can be expressed in terms of Bernoulli numbers:

$$\int_0^\infty \frac{x^{2n-1}}{\mathrm{e}^x-1}\,\mathrm{d}x = (-1)^{n-1}\frac{(2\pi)^{2n}}{4n}B_{2n} = \frac{(2\pi)^{2n}}{4n}|B_{2n}|\,. \tag{C.2.10}$$

Integrals containing the derivative of the Bose function can also be evaluated:

$$\int_0^\infty \frac{x^n\,\mathrm{e}^x}{(\mathrm{e}^x-1)^2}\,\mathrm{d}x = n!\sum_{k=1}^\infty \frac{1}{k^n} = n!\zeta(n)\,. \tag{C.2.11}$$

When the spectrum is cut off at a finite frequency, the Debye function defined as

$$D_n(x) = \frac{n}{x^n}\int_0^x \frac{t^n}{\mathrm{e}^t-1}\,\mathrm{d}t \tag{C.2.12}$$

appears. For small values of its argument

$$D_n(x) = 1 - \frac{n}{2(n+1)}x + \sum_{k=1}^\infty \frac{nB_{2k}}{(2k+n)(2k)!}x^{2k}\,, \tag{C.2.13}$$

while for large values of x

$$D_n(x) = \frac{n}{x^n}\left\{n!\zeta(n+1) - \sum_{k=1}^\infty \mathrm{e}^{-kx}\left[\frac{x^n}{k} + \frac{nx^{n-1}}{k^2} + \frac{n(n-1)x^{n-2}}{k^3} + \ldots\right]\right\}. \tag{C.2.14}$$

Through integration by parts integrals of the form

$$\mathcal{J}_n(x) = \int_0^x \frac{t^n\,\mathrm{e}^t}{(\mathrm{e}^t-1)^2}\,\mathrm{d}t\,, \tag{C.2.15}$$

which contain the derivative of the Bose function, can be rewritten as

$$\begin{aligned}\mathcal{J}_n(x) &= -\frac{x^n}{\mathrm{e}^x-1} + n\int\limits_0^x \frac{t^{n-1}}{\mathrm{e}^t-1}\,\mathrm{d}t \\ &= -\frac{x^n}{\mathrm{e}^x-1} + \frac{n}{n-1}x^{n-1}D_{n-1}(x)\,.\end{aligned} \tag{C.2.16}$$

C.2.3 Integrals Containing the Fermi Function

The definite integral over the interval $(0,\infty)$ of the product of the Fermi function $1/(\mathrm{e}^x+1)$ – which specifies the occupation of states in fermionic systems – and a power $\alpha > -1$ of x can be determined exactly:

$$\begin{aligned}\int\limits_0^\infty \frac{x^{\nu-1}}{\mathrm{e}^x+1}\,\mathrm{d}x &= \sum_{k=1}^\infty \int\limits_0^\infty x^{\nu-1}(-1)^{k-1}\mathrm{e}^{-kx}\,\mathrm{d}x \\ &= \sum_{k=1}^\infty \frac{(-1)^{k-1}}{k^\nu}\Gamma(\nu) = (1-2^{1-\nu})\Gamma(\nu)\zeta(\nu)\end{aligned} \tag{C.2.17}$$

if $\nu > 0$. For $\nu = 1$

$$\int\limits_0^\infty \frac{\mathrm{d}x}{\mathrm{e}^x+1} = \ln 2\,. \tag{C.2.18}$$

For odd powers ($\alpha = 2n-1$) the integral can be expressed with the Bernoulli numbers:

$$\int\limits_0^\infty \frac{x^{2n-1}}{\mathrm{e}^x+1}\,\mathrm{d}x = (-1)^{n-1}(1-2^{1-2n})\frac{(2\pi)^{2n}}{4n}B_{2n} = (1-2^{1-2n})\frac{(2\pi)^{2n}}{4n}|B_{2n}|\,. \tag{C.2.19}$$

Sometimes integrals containing the derivative of the Fermi function need to be evaluated:

$$\begin{aligned}\int\limits_0^\infty \frac{x^n\mathrm{e}^x}{(\mathrm{e}^x+1)^2}\,\mathrm{d}x &= \sum_{k=1}^\infty \int\limits_0^\infty x^n(-1)^{k-1}k\mathrm{e}^{-kx}\,\mathrm{d}x \\ &= n!\sum_{k=1}^\infty \frac{(-1)^{k-1}}{k^n} = n!(1-2^{1-n})\zeta(n)\,.\end{aligned} \tag{C.2.20}$$

When the chemical potential is nonzero, the *Fermi integrals*

$$F_j(\eta) = \frac{1}{j!}\int\limits_0^\infty \frac{x^j\,\mathrm{d}x}{\exp(x-\eta)+1} \tag{C.2.21}$$

appear, where the definition of $j!$ is extended to noninteger values of j via the Γ function. For $\eta = 0$

$$F_j(0) = \left(1 - 2^{-j}\right) \zeta(j+1)\,. \tag{C.2.22}$$

For negative values of η the Fermi integral can be expanded into a series as

$$F_j(\eta) = \frac{1}{j!} \sum_l \int_0^\infty x^j (-1)^{l-1} \mathrm{e}^{l\eta} \mathrm{e}^{-lx}\, \mathrm{d}x = \sum_{l=1}^{\infty} (-1)^{l-1} \frac{\mathrm{e}^{l\eta}}{l^{j+1}}\,. \tag{C.2.23}$$

For large negative values the leading-order term is

$$F_j(\eta) \approx \mathrm{e}^{\eta}\,. \tag{C.2.24}$$

When η is positive, only an asymptotic form can be obtained:

$$F_j(\eta) = \frac{\eta^{j+1}}{(j+1)!} \left[1 + \sum_{r=1}^{s} a_{2r} \eta^{-2r}\right] + R_{2s}\,, \tag{C.2.25}$$

where

$$a_{2r} = 2(j+1)j \ldots (j-2r+2)(1 - 2^{1-2r})\zeta(2r)\,, \tag{C.2.26}$$

and the remainder R_{2s} is less than

$$\frac{1}{j!}(2s+2) a_{2s+2} \eta^{j-2s-1}\,. \tag{C.2.27}$$

For $j = 1/2$ and $j = 3/2$

$$F_{1/2}(\eta) = \frac{\eta^{3/2}}{(3/2)!} \left[1 + \frac{\pi^2}{8}\eta^{-2} + \frac{7\pi^4}{640}\eta^{-4}\right] + R_4\,, \tag{C.2.28-a}$$

$$F_{3/2}(\eta) = \frac{\eta^{5/2}}{(5/2)!} \left[1 + \frac{5\pi^2}{8}\eta^{-2} - \frac{7\pi^4}{384}\eta^{-4}\right] + R_4\,. \tag{C.2.28-b}$$

C.2.4 Integrals over the Fermi Sphere

We shall repeatedly encounter sums and integrals over the Fermi sphere:

$$\begin{aligned} \frac{1}{V} \sum_{|\boldsymbol{k}|<k_\mathrm{F}} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} &= \int_{|\boldsymbol{k}|<k_\mathrm{F}} \frac{\mathrm{d}^3 k}{(2\pi)^3} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} \\ &= \frac{1}{(2\pi)^3} \int_0^{k_\mathrm{F}} k^2\, \mathrm{d}k \int_0^{\pi} \sin\theta\, \mathrm{d}\theta \int_0^{2\pi} \mathrm{d}\varphi\, \mathrm{e}^{\mathrm{i}kr\cos\theta} \\ &= \frac{k_\mathrm{F}^3}{2\pi^2} \frac{j_1(k_\mathrm{F} r)}{k_\mathrm{F} r} = \frac{k_\mathrm{F}^3}{2\pi^2} \frac{\sin k_\mathrm{F} r - k_\mathrm{F} r \cos k_\mathrm{F} r}{(k_\mathrm{F} r)^3}\,, \end{aligned} \tag{C.2.29}$$

where $j_1(x)$ is the spherical Bessel function of the first kind of order $n = 1$.

$$\frac{1}{V}\sum_{|\boldsymbol{k}'|<k_\mathrm{F}} \frac{1}{|\boldsymbol{k}-\boldsymbol{k}'|^2} = \int\limits_{|\boldsymbol{k}'|<k_\mathrm{F}} \frac{\mathrm{d}^3k'}{(2\pi)^3}\frac{1}{|\boldsymbol{k}-\boldsymbol{k}'|^2} \tag{C.2.30}$$
$$= \frac{1}{(2\pi)^3}\int\limits_0^{k_\mathrm{F}} k'^2\,\mathrm{d}k' \int\limits_0^{\pi}\sin\theta\,\mathrm{d}\theta \int\limits_0^{2\pi}\mathrm{d}\varphi \frac{1}{k^2+k'^2-2kk'\cos\theta}\,.$$

Using the substitution $\cos\theta = x$,

$$\frac{1}{V}\sum_{|\boldsymbol{k}'|<k_\mathrm{F}} \frac{1}{|\boldsymbol{k}-\boldsymbol{k}'|^2} = \frac{1}{(2\pi)^2}\int\limits_0^{k_\mathrm{F}} k'^2\,\mathrm{d}k' \int\limits_{-1}^{1}\frac{\mathrm{d}x}{k^2+k'^2-2kk'x}$$
$$= \frac{1}{(2\pi)^2}\frac{1}{k}\int\limits_0^{k_\mathrm{F}} k'\,\mathrm{d}k' \ln\left|\frac{k+k'}{k-k'}\right| \tag{C.2.31}$$
$$= \frac{k_\mathrm{F}}{2\pi^2}\left[\frac{1}{2}+\frac{k_\mathrm{F}^2-k^2}{4kk_\mathrm{F}}\ln\left|\frac{k_\mathrm{F}+k}{k_\mathrm{F}-k}\right|\right].$$

The following integral can be determined along the same lines:

$$\frac{1}{V}\sum_{|\boldsymbol{k}|<k_\mathrm{F}} \frac{1}{|\boldsymbol{k}+\boldsymbol{q}|^2-\boldsymbol{k}^2} = \frac{1}{(2\pi)^3}\int\limits_0^{k_\mathrm{F}} k^2\,\mathrm{d}k \int\limits_0^{\pi}\sin\theta\,\mathrm{d}\theta \int\limits_0^{2\pi}\mathrm{d}\varphi \frac{1}{q^2+2kq\cos\theta}$$
$$= \frac{1}{(2\pi)^2}\int\limits_0^{k_\mathrm{F}} \frac{k^2}{2kq}\,\ln\left|\frac{q^2+2kq}{q^2-2kq}\right|\,\mathrm{d}k$$
$$= \frac{k_\mathrm{F}}{8\pi^2}\left[\frac{1}{2}+\frac{4k_\mathrm{F}^2-q^2}{8qk_\mathrm{F}}\ln\left|\frac{q+2k_\mathrm{F}}{q-2k_\mathrm{F}}\right|\right]. \tag{C.2.32}$$

If (C.2.30) has to be summed with respect to the variable $\boldsymbol{k}$ on the Fermi sphere, then the introduction of the variable $\boldsymbol{k}' = \boldsymbol{k}+\boldsymbol{q}$ gives

$$\frac{1}{V^2}\sum_{|\boldsymbol{k}|<k_\mathrm{F}}\sum_{|\boldsymbol{k}'|<k_\mathrm{F}} \frac{1}{|\boldsymbol{k}-\boldsymbol{k}'|^2} = \int\limits_{|\boldsymbol{k}|<k_\mathrm{F}} \frac{\mathrm{d}^3k}{(2\pi)^3}\int\limits_{|\boldsymbol{k}'|<k_\mathrm{F}} \frac{\mathrm{d}^3k'}{(2\pi)^3}\,\frac{1}{|\boldsymbol{k}-\boldsymbol{k}'|^2}$$
$$= \iint\limits_{\substack{|\boldsymbol{k}|<k_\mathrm{F}\\ |\boldsymbol{k}+\boldsymbol{q}|<k_\mathrm{F}}} \frac{\mathrm{d}^3k}{(2\pi)^3}\frac{\mathrm{d}^3q}{(2\pi)^3}\,\frac{1}{|\boldsymbol{q}|^2}\,. \tag{C.2.33}$$

First the integration with respect to $\boldsymbol{k}$ (with $\boldsymbol{q}$ fixed) is performed. The region of integration allowed by the constraints is the intersection of two Fermi spheres with a relative displacement of $\boldsymbol{q}$ (see Fig. C.1).

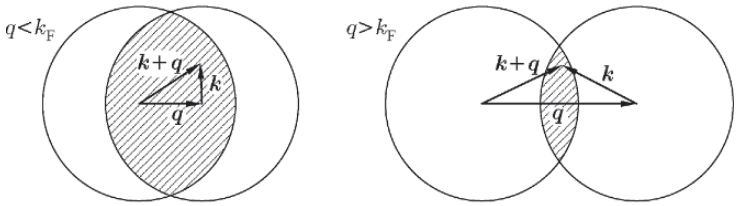


Fig. C.1. Domain of integration in the space of variables $\boldsymbol{k}$ and $\boldsymbol{k}' = \boldsymbol{k} + \boldsymbol{q}$

With $\boldsymbol{q}$ fixed, the k-integration is most easily done using cylindrical coordinates; the component along $\boldsymbol{q}$ is then denoted by k_z. Since $|\boldsymbol{q}|$ may vary between 0 and $2k_\mathrm{F}$, and as the regions $-k_\mathrm{F} < k_z < -q/2$ and $-q/2 < k_z < k_\mathrm{F} - q$ contribute equally,

$$\frac{1}{V^2} \sum_{|\boldsymbol{k}|<k_\mathrm{F}} \sum_{|\boldsymbol{k}'|<k_\mathrm{F}} \frac{1}{|\boldsymbol{k}-\boldsymbol{k}'|^2} = \frac{1}{(2\pi)^6} \int_0^{2k_\mathrm{F}} 4\pi \,\mathrm{d}q \int_{-k_\mathrm{F}}^{-q/2} \mathrm{d}k_z \, 2\pi[k_\mathrm{F}^2 - k_z^2]$$
$$= \left(\frac{k_\mathrm{F}}{2\pi}\right)^4 . \tag{C.2.34}$$

C.2.5 *d*-Dimensional Integrals

The d-dimensional integral of a function that depends only on the magnitude of the vector $\boldsymbol{k}$ is given by

$$\frac{1}{(2\pi)^d} \int f(k^2) \,\mathrm{d}^d\boldsymbol{k} = K_d \int_0^\infty f(k^2) k^{d-1} \,\mathrm{d}k \,, \tag{C.2.35}$$

where

$$K_d = \frac{1}{\pi^{d/2} 2^{d-1}} \frac{1}{\Gamma(d/2)} \,. \tag{C.2.36}$$

If the integrand depends not only on the magnitude of $\boldsymbol{k}$ but also on its component along the direction of a given vector $\boldsymbol{k}_1$,

$$\frac{1}{(2\pi)^d} \int f(k^2, \boldsymbol{k}\cdot\boldsymbol{k}_1) \,\mathrm{d}^d\boldsymbol{k} = \frac{1}{2\pi} K_{d-1} \int_0^\infty \mathrm{d}k \int_0^\pi \mathrm{d}\theta k^{d-1} (\sin\theta)^{d-2} f(k^2, k k_1 \cos\theta) \,. \tag{C.2.37}$$

C.3 Special Functions

C.3.1 The Dirac Delta Function

By definition, the Dirac delta function (or simply delta function) vanishes everywhere except for $x = 0$ where it takes an infinitely large value, in such

a way that its integral over the entire number line is unity. Some common representations are

$$\delta(x) = \lim_{\varepsilon\to 0} \frac{1}{\sqrt{\pi}\varepsilon} \mathrm{e}^{-x^2/\varepsilon^2} \tag{C.3.1-a}$$

$$= \lim_{\varepsilon\to 0} \frac{1}{\pi} \frac{\sin(x/\varepsilon)}{x} \tag{C.3.1-b}$$

$$= \lim_{\varepsilon\to 0} \frac{1}{\pi} \frac{\varepsilon}{x^2+\varepsilon^2} \tag{C.3.1-c}$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{d}k\, \mathrm{e}^{\mathrm{i}kx}. \tag{C.3.1-d}$$

The d-dimensional generalization of the last formula is

$$\delta^{(d)}(\boldsymbol{r}) = \frac{1}{(2\pi)^d} \int \mathrm{d}\boldsymbol{k}\, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\,. \tag{C.3.2}$$

The imaginary part of a function that has a simple pole close to the real axis can be expressed in terms of the Dirac delta function as

$$\frac{1}{x \pm \mathrm{i}\alpha} = \mathrm{P}\frac{1}{x} \mp \mathrm{i}\pi\delta(x)\,, \tag{C.3.3}$$

where α is infinitesimally small, and P stands for the principal value. This expression appears in the result of the integral

$$\int_0^{\infty} \mathrm{e}^{\mathrm{i}\omega t}\, \mathrm{d}t = \frac{\mathrm{i}}{\omega + \mathrm{i}\alpha} \tag{C.3.4}$$

when an infinitesimally small imaginary part is added to the variable ω to render the integral convergent at the upper limit.

The Heaviside step function can be defined as the integral of the Dirac delta function:

$$\theta(x) = \int_{-\infty}^{x} \delta(t)\, \mathrm{d}t\,, \tag{C.3.5}$$

that is

$$\delta(x) = \frac{\mathrm{d}}{\mathrm{d}x}\theta(x)\,. \tag{C.3.6}$$

An alternative representation is

$$\theta(t) = \lim_{\alpha\to 0} \frac{\mathrm{i}}{2\pi} \int_{-\infty}^{\infty} \frac{\mathrm{e}^{-\mathrm{i}\omega t}}{\omega + \mathrm{i}\alpha}\, \mathrm{d}\omega\,, \tag{C.3.7}$$

which can be obtained by taking the inverse Fourier transform of (C.3.4).

The Dirac delta function also appears in the frequently encountered formula

$$\triangle\frac{1}{r} = \boldsymbol{\nabla}^2\frac{1}{r} = -4\pi\delta(r)\,. \tag{C.3.8}$$

To demonstrate this, consider a sphere of radius R and volume V, centered at the origin, and integrate the left-hand side of the previous formula over this sphere. Using Gauss's law, the volume integral can be converted into an integral over the surface S of the sphere:

$$\int\limits_V \boldsymbol{\nabla}^2\frac{1}{r}\,\mathrm{d}\boldsymbol{r} = \int\limits_S \boldsymbol{\nabla}\frac{1}{r}\mathrm{d}\boldsymbol{S} = -\int\limits_S \frac{\mathrm{d}S}{r^2} = -4\pi\,, \tag{C.3.9}$$

which is just the integral of the right-hand side.

C.3.2 Zeta and Gamma Functions

The Riemann Zeta Function

The Riemann zeta function is defined by the sum

$$\zeta(z) = \sum_{k=1}^{\infty} k^{-z}\,, \tag{C.3.10}$$

which is convergent for $\mathrm{Re}\,z > 1$. Alternatively, it may be represented by the integral

$$\zeta(z) = \frac{1}{\Gamma(z)}\int\limits_0^{\infty}\frac{x^{z-1}}{\mathrm{e}^x-1}\mathrm{d}x = \frac{1}{\Gamma(z)}\sum_{k=1}^{\infty}\int\limits_0^{\infty}\mathrm{e}^{-kx}\int\limits_0^{\infty}x^{z-1}\,\mathrm{d}x\,. \tag{C.3.11}$$

It can be analytically continued to the entire complex plane except for the point $z = 1$.

For the first few integer values of the argument

$$\begin{aligned}
\zeta(2) &= 1+\frac{1}{2^2}+\frac{1}{3^2}+\cdots = \frac{\pi^2}{6}\,,\\
\zeta(3) &= 1+\frac{1}{2^3}+\frac{1}{3^3}+\cdots = 1.202\,,\\
\zeta(4) &= 1+\frac{1}{2^4}+\frac{1}{3^4}+\cdots = \frac{\pi^4}{90}\,,\\
\zeta(5) &= 1+\frac{1}{2^5}+\frac{1}{3^5}+\cdots = 1.037\,,\\
\zeta(6) &= 1+\frac{1}{2^6}+\frac{1}{3^6}+\cdots = \frac{\pi^6}{945}\,,
\end{aligned} \tag{C.3.12}$$

while for a couple of half-integer values

$$\zeta(3/2) = 2.612\,, \qquad \zeta(5/2) = 1.341\,. \tag{C.3.13}$$

When the sum runs over odd numbers only,

$$\sum_{k=0}^{\infty} \frac{1}{(2k+1)^n} = (1 - 2^{-n})\zeta(n)\,. \tag{C.3.14}$$

On the other hand, when the terms are added with an alternating sign, the *Dirichlet eta function* or *alternating zeta function* is obtained:

$$\eta(n) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k^n} = (1 - 2^{1-n})\zeta(n)\,. \tag{C.3.15}$$

Its integral representation can be analytically continued as

$$\eta(z) = \frac{1}{\Gamma(z)} \int_0^{\infty} \frac{x^{z-1}}{\mathrm{e}^x + 1} \mathrm{d}x = \frac{1}{\Gamma(z)} \sum_{k=1}^{\infty} \int_0^{\infty} (-1)^{k-1} \mathrm{e}^{-kx} x^{z-1}\, \mathrm{d}x\,. \tag{C.3.16}$$

For particular values of the argument

$$\eta(1) = \ln 2\,, \quad \eta(2) = \frac{\pi^2}{12}\,, \quad \eta(4) = \frac{7\pi^4}{720}\,. \tag{C.3.17}$$

The Bernoulli Numbers and their Relationship with the Zeta Function

The Bernoulli numbers B_n are defined via the series expansion

$$\frac{z}{\mathrm{e}^z - 1} = \sum_{n=0}^{\infty} B_n \frac{z^n}{n!}\,. \tag{C.3.18}$$

The first few Bernoulli numbers are

$$\begin{aligned} &B_0 = 1\,, \quad B_1 = -\frac{1}{2}\,, \quad B_2 = \frac{1}{6}\,, \quad B_{2n+1} = 0\,, \\ &B_4 = -\frac{1}{30}\,, \quad B_6 = \frac{1}{42}\,, \quad B_8 = -\frac{1}{30}\,, \quad B_{10} = \frac{5}{66}\,. \end{aligned} \tag{C.3.19}$$

For positive even (and negative odd) numbers the ζ function can also be expressed in terms of the Bernoulli numbers:

$$\zeta(2n) = (-1)^{n-1} \frac{(2\pi)^{2n}}{2(2n)!} B_{2n} \tag{C.3.20}$$

and

$$\zeta(1 - 2n) = -\frac{B_{2n}}{2n}\,, \tag{C.3.21}$$

or conversely

$$B_{2n} = \frac{(-1)^{n-1} 2(2n)!}{(2\pi)^{2n}} \zeta(2n)\,. \tag{C.3.22}$$

The Gamma Function

The gamma function (also called Euler's gamma function or the factorial function) can be defined via the recursion relation

$$\Gamma(z+1) = z\Gamma(z)\,, \tag{C.3.23}$$

with the additional condition $\Gamma(1) = 1$. In the $\mathrm{Re}\, z > 0$ region, when the condition $\mathrm{Re}\, k > 0$ is met, it can be represented by the integrals

$$\Gamma(z) = \int_0^\infty t^{z-1}\mathrm{e}^{-t}\,\mathrm{d}t = k^z \int_0^\infty t^{z-1}\mathrm{e}^{-kt}\,\mathrm{d}t\,. \tag{C.3.24}$$

It can be easily shown by integration by parts that these expressions satisfy the above recursion formula. Via analytical continuation they can be extended to the whole complex plane, where it has poles at nonpositive integers.

By iterating the recursion formula for integer values of z,

$$\Gamma(n) = (n-1)!\,. \tag{C.3.25}$$

Therefore the factorial can be defined for noninteger values of z via the gamma function. For $z = \frac{1}{2}$

$$\Gamma(\tfrac{1}{2}) = \int_0^\infty t^{-1/2}\mathrm{e}^{-t}\,\mathrm{d}t = 2\int_0^\infty \mathrm{e}^{-t^2}\,\mathrm{d}t = \sqrt{\pi}\,, \tag{C.3.26}$$

thus for half-integer values of z

$$\Gamma(n+\tfrac{1}{2}) = \frac{(2n-1)!!}{2^n}\sqrt{\pi}\,. \tag{C.3.27}$$

Consequently $\frac{1}{2}! = \Gamma(\frac{3}{2}) = \frac{1}{2}\sqrt{\pi} = 0.886$.

For large values of the argument,

$$\Gamma(x) = \sqrt{\frac{2\pi}{x}}\,x^x\,\mathrm{e}^{-x}\left\{1 + \frac{1}{12x} + \frac{1}{288x^2} + \dots\right\}. \tag{C.3.28}$$

This leads to the *Stirling formula*:

$$n! = n\Gamma(n) \approx \sqrt{2\pi n}\,n^n\,\mathrm{e}^{-n}\,. \tag{C.3.29}$$

The *digamma function* (or psi function) is defined as

$$\psi(z) = \frac{\mathrm{d}}{\mathrm{d}z}\ln\Gamma(z)\,. \tag{C.3.30}$$

Apart from negative integers, the function can be given in the form of an asymptotic series,

$$\psi(z+1) = -\gamma - \sum_{n=1}^{\infty} \left(\frac{1}{z+n} - \frac{1}{n} \right), \tag{C.3.31}$$

where γ is the Euler–Mascheroni constant (or Euler's constant):

$$\gamma = -\psi(1) = \lim_{n\to\infty} \left(\sum_{k=1}^{n} \frac{1}{k} - \ln n \right) \approx 0.577\,215\ldots . \tag{C.3.32}$$

For $z = \frac{1}{2}$

$$\psi(\tfrac{1}{2}) = -\gamma - 2\ln 2\,. \tag{C.3.33}$$

C.3.3 Bessel Functions

The Bessel functions (also called cylinder functions) of order ν are defined as solutions to the Bessel differential equation

$$z^2 \frac{\mathrm{d}^2 w(z)}{\mathrm{d}z^2} + z \frac{\mathrm{d}w(z)}{\mathrm{d}z} + (z^2 - \nu^2) w(z) = 0\,, \tag{C.3.34}$$

where ν can be any real or complex number. One particular solution can be written as a power series of z,

$$J_\nu(z) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!\Gamma(\nu+k+1)} \left(\frac{z}{2} \right)^{2k+\nu} . \tag{C.3.35}$$

Except for integer orders $\nu = n$, when

$$J_{-n}(z) = (-1)^n J_n(z)\,, \tag{C.3.36}$$

$J_\nu(z)$ and $J_{-\nu}(z)$ are linearly independent, and the general solution of the Bessel differential equation is

$$Z_\nu(z) = c_1 J_\nu(z) + c_2 J_{-\nu}(z)\,. \tag{C.3.37}$$

There are two classes of solutions to the Bessel equation for any ν with different analytic properties. The functions $J_\nu(z)$ are called Bessel functions of the first kind (or simply Bessel functions), while the Bessel functions of the second kind (more commonly called *Neumann functions* or *Weber functions*) can be written as

$$N_\nu(z) = \frac{J_\nu(z)\cos(\nu\pi) - J_{-\nu}(z)}{\sin(\nu\pi)}\,. \tag{C.3.38}$$

For integer orders n the Neumann function $N_n(z)$ is defined as the $\nu \to n$ limit of $N_\nu(z)$. The Bessel functions of the third kind, also known as *Hankel functions* are special combinations of the Bessel functions of the first and second kinds:

$$H_\nu^{(1)}(z) = J_\nu(z) + \mathrm{i}N_\nu(z)\,, \qquad H_\nu^{(2)}(z) = J_\nu(z) - \mathrm{i}N_\nu(z)\,. \tag{C.3.39}$$

For large values of $|z|$

$$J_\nu(z) \sim \sqrt{\frac{2}{\pi z}}\cos\left(z - \tfrac{1}{2}\pi\nu - \tfrac{1}{4}\pi\right), \tag{C.3.40-a}$$

$$N_\nu(z) \sim \sqrt{\frac{2}{\pi z}}\sin\left(z - \tfrac{1}{2}\pi\nu - \tfrac{1}{4}\pi\right) \tag{C.3.40-b}$$

asymptotically.

The Bessel functions of integer order have a simple integral representation:

$$J_n(z) = \frac{\mathrm{i}^{-n}}{\pi}\int_0^\pi \mathrm{e}^{\mathrm{i}z\cos\theta}\cos(n\theta)\,\mathrm{d}\theta = \frac{\mathrm{i}^{-n}}{2\pi}\int_0^{2\pi}\mathrm{e}^{\mathrm{i}z\cos\theta}\mathrm{e}^{\mathrm{i}n\theta}\,\mathrm{d}\theta\,. \tag{C.3.41}$$

Among the Bessel functions of fractional order particularly important are the Bessel functions of half-integer order. Instead of $J_{n+1/2}(z)$, it is customary to use the *spherical Bessel function of the first kind*

$$j_n(z) = \sqrt{\frac{\pi}{2z}}J_{n+1/2}(z)\,, \tag{C.3.42}$$

which satisfies the equation

$$z^2\frac{\mathrm{d}^2 j_n(z)}{\mathrm{d}z^2} + 2z\frac{\mathrm{d}j_n(z)}{\mathrm{d}z} + [z^2 - n(n+1)]j_n(z) = 0\,, \tag{C.3.43}$$

where $n = 0, \pm1, \pm2, \ldots$. It can be shown that

$$j_n(z) = z^n\left(-\frac{1}{z}\frac{\mathrm{d}}{\mathrm{d}z}\right)^n\frac{\sin z}{z}\,. \tag{C.3.44}$$

The explicit expressions for the first few spherical Bessel functions are

$$\begin{aligned} j_0(z) &= \frac{\sin z}{z}\,,\\ j_1(z) &= \frac{\sin z}{z^2} - \frac{\cos z}{z}\,,\\ j_2(z) &= \left(\frac{3}{z^3} - \frac{1}{z}\right)\sin z - \frac{3}{z^2}\cos z\,. \end{aligned} \tag{C.3.45}$$

For small values of z the leading-order contribution is

$$j_n(z) \sim \frac{z^n}{(2n+1)!!}\,, \tag{C.3.46}$$

while for large values of z

$$j_n(z) \sim \frac{1}{z}\cos\left[z - \tfrac{1}{2}(n+1)\pi\right] \sim \frac{\mathrm{i}^{-(n+1)}}{2z}\left[\mathrm{e}^{\mathrm{i}z} + (-1)^{n+1}\mathrm{e}^{-\mathrm{i}z}\right] \qquad \text{(C.3.47)}$$

asymptotically.

Just like the Bessel differential equation, (C.3.43) also has other sets of solutions, the *spherical Neumann functions* or *spherical Bessel functions of the second kind*:

$$n_n(z) = \sqrt{\frac{\pi}{2z}}N_{n+1/2}(z) = (-1)^{n+1}\sqrt{\frac{\pi}{2z}}J_{-n-1/2}(z)\,, \qquad \text{(C.3.48)}$$

and the *spherical Hankel functions* or *spherical Bessel functions of the third kind*:

$$h_n^{(1)}(z) = j_n(z) + \mathrm{i}n_n(z)\,, \qquad h_n^{(2)}(z) = j_n(z) - \mathrm{i}n_n(z)\,. \qquad \text{(C.3.49)}$$

In analogy with (C.3.44), the spherical Bessel functions of the second kind may be written as

$$n_n(z) = -z^n\left(-\frac{1}{z}\frac{\mathrm{d}}{\mathrm{d}z}\right)^n\frac{\cos z}{z}\,. \qquad \text{(C.3.50)}$$

The first few spherical Bessel functions of the second and third kinds are

$$\begin{aligned}
n_0(z) &= -\frac{\cos z}{z}\,,\\
n_1(z) &= -\frac{\cos z}{z^2} - \frac{\sin z}{z}\,, \qquad \text{(C.3.51-a)}\\
n_2(z) &= -\left(\frac{3}{z^3} - \frac{1}{z}\right)\cos z - \frac{3}{z^2}\sin z\,,\\
h_0^{(1)}(z) &= -\frac{\mathrm{i}}{z}\mathrm{e}^{\mathrm{i}z}\,,\\
h_1^{(1)}(z) &= \mathrm{e}^{\mathrm{i}z}\left(-\frac{1}{z} - \frac{\mathrm{i}}{z^2}\right), \qquad \text{(C.3.51-b)}\\
h_2^{(1)}(z) &= \mathrm{e}^{\mathrm{i}z}\left(\frac{\mathrm{i}}{z} - \frac{3}{z^2} - \frac{3\mathrm{i}}{z^3}\right).
\end{aligned}$$

For small values of z the leading-order singularity of $n_n(z)$ is of the type

$$n_n(z) \sim \frac{(2n-1)!!}{z^{n+1}}\,, \qquad \text{(C.3.52)}$$

while for large values of z

$$n_n(z) \sim \frac{1}{z}\sin\left[z - \tfrac{1}{2}(n+1)\pi\right], \qquad \text{(C.3.53-a)}$$

$$h_n^{(1)}(z) \sim \frac{1}{z}\mathrm{e}^{\mathrm{i}[z-(n+1)\pi/2]} \qquad \text{(C.3.53-b)}$$

asymptotically.

The *modified Bessel functions* of order ν satisfy the modified Bessel differential equation

$$z^2\frac{\mathrm{d}^2w(z)}{\mathrm{d}z^2}+z\frac{\mathrm{d}w(z)}{\mathrm{d}z}-(z^2+\nu^2)w(z)=0\,. \tag{C.3.54}$$

Comparison with the Bessel differential equation immediately shows that the series expansion in powers of z,

$$I_\nu(z)=\sum_{k=0}^{\infty}\frac{1}{k!\Gamma(\nu+k+1)}\left(\frac{z}{2}\right)^{\nu+2k} \tag{C.3.55}$$

is a solution. This is called *modified Bessel function of the first kind*. For integer orders $\nu=n$

$$I_n(z)=\mathrm{i}^{-n}J_n(\mathrm{i}z)\,, \tag{C.3.56}$$

while for general ν the solutions $I_\nu(z)$ and $I_{-\nu}(z)$ are linearly independent. A particular linear combination of these functions,

$$K_\nu(z)=\frac{\pi}{2}\frac{I_{-\nu}(z)-I_\nu(z)}{\sin(\nu z)} \tag{C.3.57}$$

is the *modified Bessel function of the second kind* or *Macdonald function*. It can also be obtained from the analytical continuation of the Hankel function:

$$K_\nu(z)=\tfrac{1}{2}\mathrm{i}\pi\mathrm{e}^{\mathrm{i}\pi\nu/2}H_\nu^{(1)}(\mathrm{i}z)\,. \tag{C.3.58}$$

Of special importance is the function K_0; its asymptotic form for small and large values of its argument are

$$K_0(z)=\begin{cases}-\left(\ln(\frac{1}{2}z)+\gamma\right) & z\ll 1\,,\\ \sqrt{\dfrac{\pi}{2z}}\mathrm{e}^{-z} & z\gg 1\,,\end{cases} \tag{C.3.59}$$

where $\gamma=0.57721$ is the Euler–Mascheroni constant given in (C.3.32).

C.4 Orthogonal Polynomials

C.4.1 Hermite Polynomials

The Hermite polynomials are the solutions of the differential equation

$$\frac{\mathrm{d}^2H_n(x)}{\mathrm{d}x^2}-2x\frac{\mathrm{d}H_n(x)}{\mathrm{d}x}+2nH_n(x)=0\,,\qquad n=0,1,2,\ldots\,. \tag{C.4.1}$$

A compact representation is given by Rodrigues' formula:

$$H_n(x) = (-1)^n \, \mathrm{e}^{x^2} \frac{\mathrm{d}^n}{\mathrm{d}x^n} \, \mathrm{e}^{-x^2} \tag{C.4.2}$$

They are orthogonal with respect to the weight function $w(x) = \mathrm{e}^{-x^2}$:

$$\int_{-\infty}^{\infty} \mathrm{e}^{-x^2} H_n(x) H_m(x) \, \mathrm{d}x = 2^n n! \sqrt{\pi} \delta_{nm} \, . \tag{C.4.3}$$

The explicit expressions for the first few Hermite polynomials are

$$\begin{aligned} H_0(x) &= 1 \, , \\ H_1(x) &= 2x \, , \\ H_2(x) &= 4x^2 - 1 \, , \\ H_3(x) &= 8x^3 - 12x \, , \\ H_4(x) &= 16x^4 - 48x^2 + 12 \, , \end{aligned} \tag{C.4.4}$$

while their general formula is

$$H_n(x) = \sum_{k=0}^{[n/2]} (-1)^k \frac{n!(2x)^{n-2k}}{k!(n-2k)!} \, . \tag{C.4.5}$$

C.4.2 Laguerre Polynomials

The Laguerre polynomials are the solutions of the differential equation

$$x \frac{\mathrm{d}^2 L_n(x)}{\mathrm{d}x^2} + (1-x) \frac{\mathrm{d}L_n(x)}{\mathrm{d}x} + n L_n(x) = 0 \qquad n = 0, 1, 2, \ldots \, . \tag{C.4.6}$$

Alternatively, they can be given by Rodrigues' formula:

$$L_n(x) = \frac{1}{n!} \mathrm{e}^x \frac{\mathrm{d}^n}{\mathrm{d}x^n} \left(\mathrm{e}^{-x} x^n \right) \, . \tag{C.4.7}$$

They are orthogonal with respect to the weight function $w(x) = \mathrm{e}^{-x}$:

$$\int_0^{\infty} \mathrm{e}^{-x} L_m(x) L_n(x) \, \mathrm{d}x = \delta_{m,n} \, . \tag{C.4.8}$$

The explicit expressions for the first few Laguerre polynomials are

$$\begin{aligned} L_0(x) &= 1 \, , \\ L_1(x) &= 1 - x \, , \\ L_2(x) &= 1 - 2x + \frac{1}{2} x^2 \, , \\ L_3(x) &= 1 - 3x + \frac{3}{2} x^2 - \frac{1}{6} x^3 \, . \end{aligned} \tag{C.4.9}$$

The *generalized Laguerre polynomials* are the solutions of the differential equation

$$x\frac{\mathrm{d}^2 L_n^{(\alpha)}(x)}{\mathrm{d}x^2} + (\alpha + 1 - x)\frac{\mathrm{d}L_n^{(\alpha)}(x)}{\mathrm{d}x} + nL_n^{(\alpha)}(x) = 0 \tag{C.4.10}$$

for nonnegative values of n, where α is an arbitrary complex number. Consequently, Rodrigues' formula reads

$$L_n^{(\alpha)}(x) = \frac{1}{n!}\mathrm{e}^x x^{-\alpha}\frac{\mathrm{d}^n}{\mathrm{d}x^n}\left(\mathrm{e}^{-x}x^{n+\alpha}\right). \tag{C.4.11}$$

For $\alpha > -1$ the generalized Laguerre polynomials satisfy the orthogonality relation with respect to the weight function $w(x) = \mathrm{e}^{-x}x^{\alpha}$

$$\int_0^\infty \mathrm{e}^{-x}x^{\alpha}L_n^{(\alpha)}(x)L_{n'}^{(\alpha)}(x)\,\mathrm{d}x = \frac{\Gamma(\alpha + n + 1)}{n!}\delta_{nn'}\,. \tag{C.4.12}$$

Of particular interest are the generalized Laguerre polynomials whose index α is a nonnegative integer. Their polynomial form is

$$L_n^m(x) = \sum_{k=0}^{n}(-1)^k\binom{n+m}{n-k}\frac{x^k}{k!}\,, \tag{C.4.13}$$

while their connection with ordinary Laguerre polynomials is given by

$$L_n^m(x) = (-1)^m\frac{\mathrm{d}^m}{\mathrm{d}x^m}L_{n+m}(x) \tag{C.4.14}$$

and $L_n^0(x) = L_n(x)$.

C.4.3 Legendre Polynomials

The Legendre polynomials are defined on the interval $|x| \le 1$, and satisfy the differential equation

$$(1 - x^2)\frac{\mathrm{d}^2 P_l(x)}{\mathrm{d}x^2} - 2x\frac{\mathrm{d}P_l(x)}{\mathrm{d}x} + l(l+1)P_l(x) = 0\,, \tag{C.4.15}$$

that is,

$$P_l(x) = \frac{1}{2^l l!}\frac{\mathrm{d}^l}{\mathrm{d}x^l}(x^2 - 1)^l = \frac{1}{2^l}\sum_{k=0}^{[l/2]}(-1)^k\binom{l}{k}\binom{2l-2k}{l}x^{l-2k}\,. \tag{C.4.16}$$

The Legendre polynomials are mutually orthogonal, however they are not normalized:

$$\int_{-1}^{+1} P_l(x)P_{l'}(x)\,\mathrm{d}x = \delta_{l,l'}\frac{2}{2l+1}\,. \tag{C.4.17}$$

The explicit expressions of the first few polynomials are

$$\begin{aligned}
P_0(x) &= 1\,,\\
P_1(x) &= x = \cos\theta\,,\\
P_2(x) &= \frac{1}{2}(3x^2-1) = \frac{1}{4}(3\cos 2\theta + 1)\,,\\
P_3(x) &= \frac{1}{2}(5x^3-3x) = \frac{1}{8}(5\cos 3\theta + 3\cos\theta)\,,\\
P_4(x) &= \frac{1}{8}(35x^4-30x^2+3) = \frac{1}{64}(35\cos 4\theta + 20\cos 2\theta + 9)\,.
\end{aligned} \tag{C.4.18}$$

The same equation, (C.4.15), is satisfied by the Legendre polynomials of the second kind,

$$\begin{aligned}
Q_0(x) &= \tfrac{1}{2}\ln\left(\frac{1+x}{1-x}\right),\\
Q_1(x) &= \frac{x}{2}\ln\left(\frac{1+x}{1-x}\right) - 1\,,\\
Q_2(x) &= \frac{3x^2-1}{4}\ln\left(\frac{1+x}{1-x}\right) - \frac{3}{2}x\,.
\end{aligned} \tag{C.4.19}$$

The *associated Legendre polynomials* are solutions of the equation

$$(1-x^2)\frac{\mathrm{d}^2P_l^m(x)}{\mathrm{d}x^2} - 2x\frac{\mathrm{d}P_l^m(x)}{\mathrm{d}x} + \left[l(l+1) - \frac{m^2}{1-x^2}\right]P_l^m(x) = 0\,, \tag{C.4.20}$$

that is

$$P_l^m(x) = (1-x^2)^{m/2}\frac{\mathrm{d}^m P_l(x)}{\mathrm{d}x^m} \qquad l,m = 0,1,2,\ldots \quad m \le l\,, \tag{C.4.21}$$

while the *associated Legendre polynomials of the second kind* are defined as

$$Q_l^m(x) = (1-x^2)^{m/2}\frac{\mathrm{d}^m Q_l(x)}{\mathrm{d}x^m} \qquad l,m = 0,1,2,\ldots \quad m \le l\,. \tag{C.4.22}$$

In several references an additional factor $(-1)^m$ appears in the definition.

The associated Legendre polynomials satisfy the orthogonality relation

$$\int_{-1}^{+1} P_l^m(x)P_{l'}^m(x)\,\mathrm{d}x = \delta_{l,l'}\frac{2}{2l+1}\frac{(l+m)!}{(l-m)!}\,. \tag{C.4.23}$$

The explicit expressions for the first few associated Legendre polynomials are

$$
\begin{aligned}
P_1^1(x) &= (1-x^2)^{1/2} = \sin\theta\,, \\
P_2^1(x) &= 3(1-x^2)^{1/2}x = \frac{3}{2}\sin 2\theta\,, \\
P_2^2(x) &= 3(1-x^2) = \frac{3}{2}(1-\cos 2\theta)\,, \\
P_3^1(x) &= \frac{3}{2}(1-x^2)^{1/2}(5x^2-1) = \frac{3}{8}(\sin\theta + 5\sin 3\theta)\,, \\
P_3^2(x) &= 15(1-x^2)x = \frac{15}{4}(\cos\theta - \cos 3\theta)\,, \\
P_3^3(x) &= 15(1-x^2)^{3/2} = \frac{15}{4}(3\sin\theta - \sin 3\theta)\,.
\end{aligned}
\tag{C.4.24}
$$

C.4.4 Spherical Harmonics

In physics, instead of the associated Legendre polynomials, the functions

$$
Y_l^m(\theta,\varphi) = (-1)^{\frac{m+|m|}{2}} \left(\frac{1}{2\pi}\right)^{1/2} \left[\frac{(2l+1)(l-|m|)!}{2(l+|m|)!}\right]^{1/2} P_l^{|m|}(\cos\theta)\mathrm{e}^{\mathrm{i}m\varphi}
\tag{C.4.25}
$$

are used, where $-l \le m \le l$. They are called *spherical harmonics* (or surface harmonics). Spherical harmonics satisfy the spherical harmonic differential equation, which is given by the angular part of Laplace's equation in spherical coordinates. It is immediately seen from the above form that

$$
Y_l^{-m}(\theta,\varphi) = (-1)^m Y_l^{m*}(\theta,\varphi)\,.
\tag{C.4.26}
$$

The spherical harmonics are normalized in such a way that the orthogonality relation

$$
\int_0^{2\pi}\int_0^{\pi} Y_l^{m*}(\theta,\varphi) Y_{l'}^{m'}(\theta,\varphi)\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\varphi = \delta_{ll'}\delta_{mm'}
\tag{C.4.27}
$$

and the completeness relation

$$
\begin{aligned}
\sum_{l=0}^{\infty}\sum_{m=-l}^{+l} Y_l^{m*}(\theta,\varphi) Y_l^m(\theta',\varphi') &= \frac{\delta(\theta-\theta')\delta(\varphi-\varphi')}{\sin\theta} \\
&= \delta(\cos\theta - \cos\theta')\delta(\varphi-\varphi') \equiv \delta(\Omega-\Omega')
\end{aligned}
\tag{C.4.28}
$$

are both satisfied.

According to the addition theorem for spherical harmonics, the product of two spherical harmonics can be written as the linear combination of spherical harmonics:

$$
\begin{aligned}
Y_{l_1}^{m_1}(\theta,\varphi) Y_{l_2}^{m_2}(\theta,\varphi) = &\sum_{l=|l_1-l_2|}^{l_1+l_2}\sum_{m=-l}^{l} \sqrt{\frac{(2l_1+1)(2l_2+1)}{4\pi(2l+1)}} \\
&\times (l_1 l_2 m_1 m_2 | lm)(l_1 l_2 00 | l0) Y_l^m(\theta,\varphi)\,.
\end{aligned}
\tag{C.4.29}
$$

Since only the $m = m_1 + m_2$ terms give nonvanishing contribution, this can be rewritten in the form

$$Y_{l_1}^{m_1}(\theta,\varphi)Y_{l_2}^{m_2}(\theta,\varphi) = \sum_l \sqrt{\frac{(2l_1+1)(2l_2+1)(2l+1)}{4\pi}} \times \begin{pmatrix} l_1 & l_2 & l \\ m_1 & m_2 & -(m_1+m_2) \end{pmatrix} \begin{pmatrix} l_1 & l_2 & l \\ 0 & 0 & 0 \end{pmatrix} Y_l^{-(m_1+m_2)^*}(\theta,\varphi)\,, \tag{C.4.30}$$

where the square-root term on the right-hand side is a Clebsch–Gordan coefficient, while the parenthesized terms are the Wigner $3j$ symbols. Both will be discussed in detail in Appendix F.

This formula and the orthogonality relation for spherical harmonics imply that the integral of the product of three spherical harmonics is

$$\int \mathrm{d}\Omega\, Y_{l_1}^{m_1}(\theta,\varphi)Y_{l_2}^{m_2}(\theta,\varphi)Y_{l_3}^{m_3}(\theta,\varphi) = \sqrt{\frac{(2l_1+1)(2l_2+1)(2l_3+1)}{4\pi}} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix}. \tag{C.4.31}$$

The explicit expressions for the first few spherical harmonics are

$$\begin{aligned}
Y_0^0 &= \left(\frac{1}{4\pi}\right)^{1/2}, \\
Y_1^0 &= \left(\frac{3}{4\pi}\right)^{1/2} \cos\theta\,, \\
Y_1^{\pm1} &= \mp\left(\frac{3}{8\pi}\right)^{1/2} \sin\theta \mathrm{e}^{\pm\mathrm{i}\varphi}\,, \\
Y_2^0 &= \left(\frac{5}{16\pi}\right)^{1/2} (3\cos^2\theta - 1)\,, \\
Y_2^{\pm1} &= \mp\left(\frac{15}{8\pi}\right)^{1/2} \sin\theta\cos\theta \mathrm{e}^{\pm\mathrm{i}\varphi}\,, \\
Y_2^{\pm2} &= \left(\frac{15}{32\pi}\right)^{1/2} \sin^2\theta \mathrm{e}^{\pm2\mathrm{i}\varphi}\,, \\
Y_3^0 &= \left(\frac{7}{16\pi}\right)^{1/2} (5\cos^3\theta - 3\cos\theta)\,, \\
Y_3^{\pm1} &= \mp\left(\frac{21}{64\pi}\right)^{1/2} (5\cos^2\theta - 1)\sin\theta \mathrm{e}^{\pm\mathrm{i}\varphi}\,, \\
Y_3^{\pm2} &= \left(\frac{105}{32\pi}\right)^{1/2} \sin^2\theta\cos\theta \mathrm{e}^{\pm2\mathrm{i}\varphi}\,, \\
Y_3^{\pm3} &= \mp\left(\frac{35}{64\pi}\right)^{1/2} \sin^3\theta \mathrm{e}^{\pm3\mathrm{i}\varphi}\,.
\end{aligned} \tag{C.4.32}$$

C.4.5 Expansion in Spherical Harmonics

It follows from the orthogonality and completeness relations that any function $f(\theta, \varphi)$ that depends only on the polar angles can be expanded in spherical harmonics:

$$f(\theta, \varphi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} c_{lm} Y_l^m(\theta, \varphi) \,, \tag{C.4.33}$$

where

$$c_{lm} = \int Y_l^{m*}(\theta, \varphi) f(\theta, \varphi) \, \mathrm{d}\Omega \,. \tag{C.4.34}$$

Another expansion is used for functions that depend on the angle between two vectors specified by spherical coordinates. We shall denote the angle between the directions specified by the polar angles θ, φ and θ', φ' by ζ. They are related by

$$\cos\zeta = \cos\theta \cos\theta' + \sin\theta \sin\theta' \cos(\varphi - \varphi') \,. \tag{C.4.35}$$

In the expansion in associated Legendre polynomials or spherical harmonics

$$\begin{aligned} P_l(\cos\zeta) &= P_l(\cos\theta) P_l(\cos\theta') \\ &\quad + 2 \sum_{m=1}^{l} \frac{(l-m)!}{(l+m)!} P_l^m(\cos\theta) P_l^m(\cos\theta') \cos m(\varphi - \varphi') \\ &= \frac{4\pi}{2l+1} \sum_{m=-l}^{+l} Y_l^{m*}(\theta, \varphi) Y_l^m(\theta', \varphi') \,. \end{aligned} \tag{C.4.36}$$

By applying the relation

$$\mathrm{e}^{\mathrm{i}xy} = \sum_{l=0}^{\infty} (2l+1) \mathrm{i}^l j_l(y) P_l(x) \tag{C.4.37}$$

to the expansion of a plane wave,

$$\begin{aligned} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} &= \sum_{l=0}^{\infty} (2l+1) \mathrm{i}^l j_l(kr) P_l(\cos\theta) \\ &= 4\pi \sum_{l=0}^{\infty} \sum_{m=-l}^{+l} \mathrm{i}^l j_l(kr) Y_l^{m*}(\theta_k, \varphi_k) Y_l^m(\theta_r, \varphi_r) \,, \end{aligned} \tag{C.4.38}$$

where (θ_k, φ_k), and (θ_r, φ_r) are the polar angles of the vectors $\boldsymbol{k}$ and $\boldsymbol{r}$.

The so-called multipole expansion of the Coulomb potential is

$$\begin{aligned} \frac{1}{|\boldsymbol{r} - \boldsymbol{r}'|} &= \sum_{l=0}^{\infty} \frac{r_<^l}{r_>^{l+1}} P_l(\cos\zeta) \\ &= \sum_{l=0}^{\infty} \frac{r_<^l}{r_>^{l+1}} \frac{4\pi}{2l+1} \sum_{m=-l}^{+l} Y_l^{m*}(\theta, \varphi) Y_l^m(\theta', \varphi') \,, \end{aligned} \tag{C.4.39}$$

where $r_<$ ($r_>$) denotes the length of the shorter (longer) vector of $\boldsymbol{r}$ and $\boldsymbol{r}'$.

As will be shown in Chapter 16 (Volume 2), the Green function of free electrons can be written as

$$G(\boldsymbol{r},\boldsymbol{r}') = -\frac{m_\mathrm{e}}{2\pi\hbar^2}\frac{\mathrm{e}^{\mathrm{i}k|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|}\,, \tag{C.4.40}$$

since

$$\left(\boldsymbol{\nabla}^2+k^2\right)\frac{\mathrm{e}^{\mathrm{i}k|\boldsymbol{r}|}}{|\boldsymbol{r}|} = -4\pi\delta(\boldsymbol{r})\,, \quad \left(\boldsymbol{\nabla}^2+k^2\right)\frac{\cos k|\boldsymbol{r}|}{|\boldsymbol{r}|} = -4\pi\delta(\boldsymbol{r})\,. \tag{C.4.41}$$

The Green function, as well as its real and imaginary parts can be expressed in terms of the spherical harmonics

$$\begin{aligned}\frac{\mathrm{e}^{\mathrm{i}k|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|} &= \mathrm{i}k\sum_{l=0}^{\infty}(2l+1)j_l(kr_<)h_l^{(1)}(kr_>)P_l(\cos\zeta)\\ &= 4\pi\,\mathrm{i}k\sum_{l=0}^{\infty}\sum_{m=-l}^{+l} j_l(kr_<)h_l^{(1)}(kr_>)Y_l^{m*}(\theta,\varphi)Y_l^m(\theta',\varphi')\,,\end{aligned} \tag{C.4.42}$$

$$\begin{aligned}\frac{\sin(k|\boldsymbol{r}-\boldsymbol{r}'|)}{|\boldsymbol{r}-\boldsymbol{r}'|} &= k\sum_{l=0}^{\infty}(2l+1)j_l(kr_<)j_l(kr_>)P_l(\cos\zeta) \qquad \text{(C.4.43)}\\ &= 4\pi k\sum_{l=0}^{\infty}j_l(kr_<)j_l(kr_>)\sum_{m=-l}^{+l}Y_l^{m*}(\theta,\varphi)Y_l^m(\theta',\varphi')\,,\\ \frac{\cos(k|\boldsymbol{r}-\boldsymbol{r}'|)}{|\boldsymbol{r}-\boldsymbol{r}'|} &= -k\sum_{l=0}^{\infty}(2l+1)j_l(kr_<)n_l(kr_>)P_l(\cos\zeta) \qquad \text{(C.4.44)}\\ &= -4\pi k\sum_{l=0}^{\infty}j_l(kr_<)n_l(kr_>)\sum_{m=-l}^{+l}Y_l^{m*}(\theta,\varphi)Y_l^m(\theta',\varphi')\end{aligned}$$

where ζ is the angle between the directions of $\boldsymbol{r}$ and $\boldsymbol{r}'$, while (θ,φ) and (θ',φ') are the polar angles of $\boldsymbol{r}$ and $\boldsymbol{r}'$.

It follows from the orthogonality relation for associated Legendre polynomials that if the two functions

$$A_l(\theta,\varphi) = a_0P_l(\cos\theta) + \sum_{m=1}^{l}\left(a_m\cos m\varphi + b_m\sin m\varphi\right)P_l^m(\cos\theta)\,, \tag{C.4.45}$$

$$B_{l'}(\theta,\varphi) = \alpha_0P_{l'}(\cos\theta) + \sum_{m=1}^{l'}\left(\alpha_m\cos m\varphi + \beta_m\sin m\varphi\right)P_{l'}^m(\cos\theta) \tag{C.4.46}$$

are introduced with arbitrary coefficients, the following formulas are valid:

$$\int_0^{2\pi} \mathrm{d}\varphi \int_0^{\pi} A_l(\theta,\varphi) B_{l'}(\theta,\varphi) \sin\theta \,\mathrm{d}\theta = 0 \tag{C.4.47}$$

if $l \neq l'$, and

$$\int_0^{2\pi} \mathrm{d}\varphi \int_0^{\pi} A_l(\theta,\varphi) P_l(\cos\zeta) \sin\theta \,\mathrm{d}\theta = \frac{4\pi}{2l+1} A_l(\theta',\varphi') \,. \tag{C.4.48}$$

The general solution of the free Schrödinger equation in spherical coordinates can be written as

$$\psi(r',\theta',\varphi') = \sum_{l,m} f_l(r') Y_l^m(\theta',\varphi') \,, \tag{C.4.49}$$

where the $f_l(r')$ are spherical Bessel functions of the first or second kind. The same function can be expanded around a different origin (displaced by $\boldsymbol{r}$ with respect to the first), using the spherical coordinates r'', θ'', φ'' of the vector $\boldsymbol{r}'' = \boldsymbol{r} + \boldsymbol{r}'$. The obtained functions $f_l(r'')$ are related to the $f_l(r')$ via the Wigner $3j$ symbols:

$$f_l(r'') Y_l^m(\theta'',\varphi'') = \sum_{l',l'',m'} \mathrm{i}^{l'+l''-l} (-1)^m (2l'+1)(2l''+1) \begin{pmatrix} l & l' & l'' \\ 0 & 0 & 0 \end{pmatrix} \times \begin{pmatrix} l & l' & l'' \\ m & -m' & m'-m \end{pmatrix} j_{l''}(r) Y_{l''}^{m-m'}(\theta,\varphi) f_{l'}(r') Y_{l''}^{m'}(\theta',\varphi') \,, \tag{C.4.50}$$

where r, θ, φ are the spherical coordinates of $\boldsymbol{r}$.

References

1. G. B. Arfken and H.-J. Weber, *Mathematical Methods for Physicists*, Sixth Edition, Elsevier/Academic Press, Amsterdam (2005).
2. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Edited by M. Abramowitz and I. A. Stegun, 10th printing, Dover Publications, Inc., New York (1972). It is available online at www.math.sfu.ca/∼cbm/aands.
3. I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series and Products*, Sixth Edition, Edited by A. Jeffrey and D. Zwillinger, Academic Press, Inc., Boston (2000).

4. A. P. Prudnikov, Yu. A. Brychkov and O. I. Marichev, *Integrals and Series*, Vol. 1. *Elementary Functions*, Gordon and Breach Science Publishers, New York (1986).
5. A. P. Prudnikov, Yu. A. Brychkov and O. I. Marichev, *Integrals and Series*, Vol. 2. *Special Functions*, Gordon and Breach Science Publishers, New York (1990).

D

Fundamentals of Group Theory

As mentioned in Chapter 6, crystalline materials occupy a special place among solids not in the least because the symmetries of their structure manifest themselves in macroscopic properties, which highly facilitates the theoretical description of their behavior. Symmetry operations form a group in the mathematical sense, therefore statements of general validity can often be based on group-theoretical considerations. In this appendix we shall review some basic notions and relationships of group theory, then present the irreducible representations of the O_h group, and finally list the group theoretical theorems related to the quantum mechanical eigenvalue problem.

D.1 Basic Notions of Group Theory

There exists a whole wealth of group theoretical textbooks written by physicists for physicists. A few of them are listed among the references. These contain plenty of solid-state physics applications. Therefore below we shall discuss only a few basic notions and theorems.

D.1.1 Definition of Groups

The set of elements $R, S, T, \ldots$ are said to form a group $\mathcal{G}$ if a binary operation called multiplication is defined among group elements, which satisfies the following requirements:

1. For any two elements R and S the product, denoted as RS, is an element T of the group:
$$RS = T \qquad (R, S, T \in \mathcal{G}) . \tag{D.1.1}$$
The operation of group multiplication does not need to be commutative: in general the order of the elements is not immaterial. A group is called commutative or Abelian if group multiplication satisfies the commutative law, i.e., $RS = SR$ for any two elements R and S.

2. Multiplication satisfies the associative law, i.e., for any elements R, S, and T
$$R(S\,T) = (R\,S)T\,, \tag{D.1.2}$$
where, naturally, multiplication within the parentheses has to be performed first.
3. There exists an identity element (also called unit element) E in the set. When any element R of the group is multiplied by E, it remains unchanged:
$$R\,E = E\,R = R\,. \tag{D.1.3}$$
4. For any element R of the group there exists a unique element S in the group such that
$$R\,S = S\,R = E\,. \tag{D.1.4}$$
This element S is called the inverse of R, and is denoted by $S \equiv R^{-1}$.

The number g of elements in the set is called the order of the group. The number of elements can be either finite or infinite; these correspond to finite and infinite groups. In solid-state physics applications we deal mostly with finite groups.

When the group is finite, the elements in the sequence R, R^2, R^3, ... repeat themselves from a certain point onward. Consequently one of these elements must be the identity element. The order of an element R in the group is the least positive integer n such that $R^n = E$.

A subset $\mathcal{G}'$ of the group elements is called a subgroup of $\mathcal{G}$ if it satisfies in itself the previous requirements that define a group – i.e., it contains the identity element, for every element in $\mathcal{G}'$ its inverse is also in $\mathcal{G}'$, and the product of any two elements in $\mathcal{G}'$ is also in $\mathcal{G}'$.

A subgroup $\mathcal{G}'$ of the group $\mathcal{G}$ is an *invariant subgroup* or *normal subgroup* if $R\,S\,R^{-1} \in \mathcal{G}'$ for every $R \in \mathcal{G}$ and every $S \in \mathcal{G}'$.

If $R_1, S_1, T_1, \ldots$ are the elements of a group $\mathcal{G}_1$ of order g_1, and R_2, S_2, $T_2, \ldots$ are the elements of a group $\mathcal{G}_2$ of order g_2, then the *direct product* (or *Kronecker product*) of the two groups, denoted by $\mathcal{G}_1 \otimes \mathcal{G}_2$ is the group made up of the elements (R_1, R_2), $(R_1, S_2), \ldots$, (S_1, R_2), $(S_1, S_2), \ldots$, with the group multiplication defined as
$$(R_1, S_2)(T_1, U_2) = (R_1 T_1, S_2 U_2)\,. \tag{D.1.5}$$
This group is of order $g_1\,g_2$.

If $\mathcal{G}_1$ and $\mathcal{G}_2$ are such subgroups of $\mathcal{G}$ that 1.) the elements in $\mathcal{G}_1$ commute with the elements in $\mathcal{G}_2$; 2.) they have a single element in common, the identity element; and 3.) each element in $\mathcal{G}$ can be written as the product of an element in $\mathcal{G}_1$ and an element in $\mathcal{G}_2$, then the group $\mathcal{G}$ is isomorphic to the direct product group $\mathcal{G}_1 \otimes \mathcal{G}_2$.

The group $\mathcal{G}$ is the semidirect product of its subgroups $\mathcal{G}_1$ and $\mathcal{G}_2$ ($\mathcal{G} = \mathcal{G}_1 \rtimes \mathcal{G}_2$), if 1.) the elements in $\mathcal{G}_1$ do not commute with the elements in $\mathcal{G}_2$

but $\mathcal{G}_1$ is an invariant subgroup; 2.) they have a single element in common, the identity element; and 3.) each element in $\mathcal{G}$ can be written as the product of an element in $\mathcal{G}_1$ and an element in $\mathcal{G}_2$.

D.1.2 Conjugate Elements and Conjugacy Classes

The group element S is conjugate to R if there exists at least one element U in the group such that $S = URU^{-1}$. This property is mutual, since along with U, its inverse U^{-1} is also an element of the group, consequently whenever the relation $S = URU^{-1}$ holds, so does $R = U^{-1}S(U^{-1})^{-1}$. Thus S and R are said to be conjugate.

The conjugate property is transitive, that is if S and R are conjugate ($S = URU^{-1}$) and so are T and S ($T = VSV^{-1}$) then T and R are conjugate, too:

$$T = VSV^{-1} = VURU^{-1}V^{-1} = (VU)R(VU)^{-1}\,. \tag{D.1.6}$$

Therefore the elements of the group can be unambiguously divided into classes of conjugate elements, *conjugacy classes*.

The multiplication rule $RS = SR$ of Abelian groups implies that for any R and S the relation $R = S^{-1}RS$ holds. Then each element is only conjugate to itself, and so forms a separate conjugacy class.

D.1.3 Representations and Characters

Consider a d-dimensional vector space spanned by the linearly independent vectors $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, …, $\boldsymbol{e}_d$. Now define a nonsingular linear transformation O on this space such that each point

$$\boldsymbol{r} = \sum_{i=1}^{d} a_i \boldsymbol{e}_i \tag{D.1.7}$$

of the space is taken into another point

$$\boldsymbol{r}' = O\boldsymbol{r} = \sum_{i=1}^{d} b_i \boldsymbol{e}_i\,. \tag{D.1.8}$$

Expressing the transforms of vectors $\boldsymbol{e}_i$ in terms of the original vectors, the matrix D defined through

$$O\boldsymbol{e}_i = \sum_j D_{ji}\boldsymbol{e}_j \tag{D.1.9}$$

specifies the mapping between the components of $\boldsymbol{r}$ and $\boldsymbol{r}'$:

$$b_i = \sum_j D_{ij} a_j\,. \tag{D.1.10}$$

If the transformation between $\boldsymbol{r}$ and $\boldsymbol{r}'$ is a one-to-one correspondence, then the inverse of the transformation O exists,

$$\boldsymbol{r} = O^{-1}\boldsymbol{r}' \,, \tag{D.1.11}$$

and this mapping can be given by the inverse of the matrix D:

$$a_i = \sum_j \left(D^{-1}\right)_{ij} b_j \,. \tag{D.1.12}$$

Consider a group $\mathcal{G}$, and associate a linear operator $O(R)$ and the corresponding matrix $D_{ij}(R)$ with each element R. If the same multiplication rule applies to the operators as to the group elements, i.e., for $RS = T$

$$O(R)O(S) = O(T) \,, \tag{D.1.13}$$

then the analogous relation

$$\mathsf{D}(R)\mathsf{D}(S) = \mathsf{D}(T) \tag{D.1.14}$$

applies to the matrices, too. In component form

$$\sum_k D_{jk}(R)D_{ki}(S) = D_{ji}(T) \,, \tag{D.1.15}$$

as

$$\begin{aligned} O(R)O(S)\boldsymbol{e}_i &= O(R)\sum_k D_{ki}(S)\boldsymbol{e}_k \qquad\qquad (\text{D.1.16}) \\ &= \sum_k \sum_j D_{ki}(S)D_{jk}(R)\boldsymbol{e}_j = \sum_j \left[\mathsf{D}(R)\mathsf{D}(S)\right]_{ji}\boldsymbol{e}_j \,. \end{aligned}$$

The identity element E is associated with the identity operator and the unit matrix, while the inverse of the element R is associated with the inverse of the operator $O(R)$ and the inverse of the matrix $\mathsf{D}(R)$. These matrices provide a *representation* of the group. The representation is *faithful* if the mapping is not only homomorphic but also isomorphic, i.e., there exists a one-to-one correspondence.

Two representations of a group are equivalent if there exists a unitary matrix U $\left(\mathsf{U}^\dagger = \mathsf{U}^{-1}\right)$ that transforms the matrices $\mathsf{D}^{(1)}(R)$ and $\mathsf{D}^{(2)}(R)$ belonging to the two representations into each other for each element of the group:

$$\mathsf{D}^{(2)}(R) = \mathsf{U}\mathsf{D}^{(1)}(R)\mathsf{U}^{-1} \,. \tag{D.1.17}$$

For every representation of a finite group there exists an equivalent representation in which the matrices are unitary. In what follows we shall only deal with unitary representations.

Being an invariant expression of the matrix elements, the diagonal sum (trace) of the matrices in the representation of the group plays an important

role in the characterization of the representations. The trace of matrix $\mathsf{D}(R)$ that represents the element R is called the *character* of the representation associated with the group element:

$$\chi(R) = \operatorname{Tr} \mathsf{D}(R) = \sum_i D_{ii}(R)\,. \tag{D.1.18}$$

In any representation the characters of mutually conjugate elements (elements of the same conjugacy class) are identical, since if $S = T\,R\,T^{-1}$ then

$$\chi(S) = \operatorname{Tr}\left(\mathsf{D}(T)\mathsf{D}(R)\mathsf{D}(T^{-1})\right) = \operatorname{Tr} \mathsf{D}(R) = \chi(R)\,. \tag{D.1.19}$$

Equivalent representations are associated with identical characters.

D.1.4 Reducible and Irreducible Representations

The matrices representing the group elements can be separately diagonalized by a unitary transformation $\mathsf{D}'(R) = \mathsf{U}\,\mathsf{D}(R)\mathsf{U}^{-1}$, however different matrices U are usually required for different elements. A representation is called reducible if there exists a matrix U that simultaneously block diagonalizes the matrices associated with each group element. If there exist no such unitary transformations that would lead to such a block-diagonal form, the representation is said to be irreducible.

Determining whether a representation is reducible or irreducible, finding the irreducible representations, and the reduction of a reducible representation rank among the most important applications of group theory in solid-state physics. These tasks are facilitated by the following theorems (presented without proof).

The number r of irreducible representations $D^{(1)}$, $D^{(2)}$, ..., $D^{(r)}$ of a finite group is equal to the number of conjugacy classes in the group.

The dimensions d_μ of irreducible representations and the number g of group elements (the order of the group) are related by

$$\sum_{\mu=1}^{r} d_\mu^2 = g\,. \tag{D.1.20}$$

The matrix elements of the matrices $\mathsf{D}^{(\mu)}$ and $\mathsf{D}^{(\nu)}$ belonging to the unitary irreducible representations $D^{(\mu)}$ and $D^{(\nu)}$ satisfy the orthogonality relation

$$\sum_{R\in\mathcal{G}} {D_{ij}^{(\mu)}}^*(R) D_{kl}^{(\nu)}(R) = \frac{g}{d_\mu}\delta_{ik}\delta_{jl}\delta_{\mu\nu}\,, \tag{D.1.21}$$

where d_μ is the dimension of the μth irreducible representation.

This implies the following formula for the characters of unitary irreducible representations:

$$\sum_{R\in\mathcal{G}} {\chi^{(\mu)}}^*(R)\chi^{(\nu)}(R) = g\delta_{\mu\nu}\,. \tag{D.1.22}$$

Since the character is the same for elements in the same class, the sum over the group elements can be replaced by a sum over the conjugacy classes. Denoting the number of elements in the ith class by g_i, and the character in the μth irreducible representation of the elements in the ith class by $\chi_i^{(\mu)}$,

$$\sum_{i=1}^{r} g_i \chi_i^{(\mu)^*} \chi_i^{(\nu)} = g\delta_{\mu\nu} \,. \tag{D.1.23}$$

The characters of the irreducible representations satisfy another orthogonality relation,

$$\sum_{\mu=1}^{r} \chi_i^{(\mu)^*} \chi_j^{(\mu)} = \frac{g}{g_i}\delta_{ij} \,. \tag{D.1.24}$$

For the point groups most often encountered in solid-state physics, the characters of the irreducible representations can be looked up in a various references. The present appendix presents one such example: the character table of the group O_h is given in Table D.1 on page 648.

D.1.5 The Reduction of Reducible Representations

According to (D.1.22), the characters of irreducible representations must satisfy the relation

$$\frac{1}{g}\sum_{R\in\mathcal{G}} |\chi^{(\mu)}(R)|^2 = 1 \,. \tag{D.1.25}$$

This necessary condition is at the same time the sufficient condition for the irreducibility of the representation. If the characters $\chi(R)$ of a representation do not satisfy the relation

$$\frac{1}{g}\sum_{R\in\mathcal{G}} |\chi(R)|^2 = 1 \tag{D.1.26}$$

then the representation is reducible.

The matrix D of any reducible representation D can be block diagonalized with a suitably chosen unitary transformation U in such a way that the matrices of the irreducible representations appear in the individual blocks. The reducible representation is then said to have been decomposed into the direct sum of irreducible representations:

$$\begin{aligned} D(R) &= D^{(1)}(R) \oplus D^{(1)}(R) \oplus \cdots \oplus D^{(2)}(R) \oplus D^{(2)}(R) \oplus \ldots \\ &= \sum_{\mu=1}^{r} n_\mu D^{(\mu)}(R) \,. \end{aligned} \tag{D.1.27}$$

It follows from this decomposition that the characters of the reducible representation can be expressed with the characters of the irreducible representations:

$$\chi(R) = \sum_{\mu=1}^{r} n_\mu \chi^{(\mu)}(R)\,. \tag{D.1.28}$$

Since their character is the same, for any element in the ith conjugacy class

$$\chi_i = \sum_{\mu=1}^{r} n_\mu \chi_i^{(\mu)}\,. \tag{D.1.29}$$

The number n_μ, which specifies how many times the irreducible representation $D^{(\mu)}$ appears in the reduction formula, can be determined from the relation

$$n_\mu = \frac{1}{g}\sum_{R\in\mathcal{G}} \chi^{(\mu)^*}(R)\chi(R) = \frac{1}{g}\sum_{i=1}^{r} g_i \chi_i^{(\mu)^*}\chi_i\,. \tag{D.1.30}$$

Based on this formula, the reduction of an arbitrary representation is straightforward when its character and the characters of the irreducible representations are known.

D.1.6 Compatibility Condition

In applications, where certain symmetries are broken by the perturbing potential the relationship between the irreducible representations of a group and of one of its subgroups may become particularly important. Consider a group $\mathcal{G}$ and a subgroup $\mathcal{G}_1$ thereof. An irreducible representation of the group $\mathcal{G}$ may be reducible on the elements of the subgroup $\mathcal{G}_1$. Usually only a small number of the irreducible representations of the subgroup $\mathcal{G}_1$ appear in the reduction formula. These irreducible representations of the subgroup are said to be *compatible* with the corresponding irreducible representations of the group $\mathcal{G}$.

D.1.7 Basis Functions of the Representations

The representation matrices were defined to act on a linear space. However, no significance has been attached to the basis vectors of the space up to now. Physical applications are especially interesting when group elements act on a function space rather than an abstract linear space. Then the basis functions of irreducible representations exhibit special properties.

To define the action of the group elements on a function space, consider the coordinate transformations that are of particular importance in solid-state physics. The coordinate transformation R, which moves point $\boldsymbol{r}$ into $\boldsymbol{r}' = R\boldsymbol{r}$, is associated with a linear operator $O(R)$ that acts on the space of functions and takes the function $\psi(\boldsymbol{r})$ into $\psi'(\boldsymbol{r})$

$$O(R)\psi(\boldsymbol{r}) = \psi'(\boldsymbol{r})\,. \tag{D.1.31}$$

The transformation is defined by the requirement that the transformed function take the same value at $\boldsymbol{r}'$ as the original function at $\boldsymbol{r}$, that is

$$\psi'(\boldsymbol{r}') = O(R)\psi(\boldsymbol{r}') = \psi(\boldsymbol{r})\,, \tag{D.1.32}$$

which implies

$$O(R)\psi(\boldsymbol{r}) = \psi(R^{-1}\boldsymbol{r})\,. \tag{D.1.33}$$

The operators $O(R)$ form a group with the same multiplication rule as the group of coordinate transformations. Since the mapping of the group of coordinate transformations on the group of these operators is homomorphic, any representation of the group of operators is also a representation of the group of coordinate transformations.

Now select an arbitrary function $\psi(\boldsymbol{r})$, and act on it with the operators $O(R)$ that belong to a group $\mathcal{G}$ of coordinate transformations. When the group is of order g, the number of functions obtained this way is also g. However, they are not necessarily linearly independent. In the space of transformed functions a linearly independent set of basis functions $(\phi_1, \phi_2, \ldots, \phi_r)$ can be chosen. Expanding $\psi(\boldsymbol{r})$ in terms of these,

$$\psi(\boldsymbol{r}) = \sum_{i=1}^{r} a_i \phi_i(\boldsymbol{r})\,. \tag{D.1.34}$$

Applying the transformation $O(R)$ on the basis functions,

$$O(R)\phi_i(\boldsymbol{r}) = \phi_i(R^{-1}\boldsymbol{r}) = \sum_{j=1}^{r} D_{ji}(R)\phi_j(\boldsymbol{r})\,. \tag{D.1.35}$$

The matrix $\mathsf{D}(R)$ in the previous formula provides a representation of the group on the space spanned by the functions $\phi_i(\boldsymbol{r})$. When the transformed functions $\psi(R^{-1}\boldsymbol{r})$ are also expanded using this basis, the coefficients in the formula

$$\psi(R^{-1}\boldsymbol{r}) = \sum_{i=1}^{r} b_i(R)\phi_i(\boldsymbol{r}) \tag{D.1.36}$$

and in the expansion of $\psi(\boldsymbol{r})$ are found to be related by

$$b_i(R) = \sum_{j=1}^{r} D_{ij}(R) a_j\,. \tag{D.1.37}$$

The representation generated above is not necessarily irreducible, therefore an additional reduction procedure may be required. It may then be established that any function $\psi(\boldsymbol{r})$ can be expanded into suitably chosen basis functions of the irreducible representations,

$$\psi(\boldsymbol{r}) = \sum_{\mu} \sum_{i=1}^{d_\mu} a_i^{(\mu)} \phi_i^{(\mu)}(\boldsymbol{r})\,. \tag{D.1.38}$$

There is a certain arbitrariness in the choice of the basis functions of the irreducible representations since they are not unambiguously determined. However, regardless of the particular choice of the basis functions, some important relationships among them are always valid.

If a group $\mathcal{G}$ of coordinate transformations is represented irreducibly on the function spaces $\phi_1^{(\mu)}(\boldsymbol{r})$, $\phi_2^{(\mu)}(\boldsymbol{r})$, ... and $\phi_1^{(\nu)}(\boldsymbol{r})$, $\phi_2^{(\nu)}(\boldsymbol{r})$, ... by the unitary irreducible representations $D^{(\mu)}(R)$ and $D^{(\nu)}(R)$, then

$$\left(\phi_i^{(\mu)}, \phi_j^{(\nu)}\right) = \int \mathrm{d}\boldsymbol{r}\, \phi_i^{(\mu)*}(\boldsymbol{r})\phi_j^{(\nu)}(\boldsymbol{r}) = A^{(\mu)}\delta_{\mu\nu}\delta_{ij}\,, \tag{D.1.39}$$

that is, two basis functions with different labels are orthogonal. The product of two basis functions with the same label depends only on the choice of the representation but not on the choice of the basis function.

It is not too difficult to construct an operator

$$P_{ij}^{(\mu)} = \frac{d_\mu}{g}\sum_{R\in\mathcal{G}} D_{ij}^{(\mu)*}(R)O(R) \tag{D.1.40}$$

such that

$$P_{ij}^{(\mu)}\phi_k^{(\nu)}(\boldsymbol{r}) = \delta_{\mu\nu}\delta_{jk}\phi_i^{(\mu)}(\boldsymbol{r})\,. \tag{D.1.41}$$

The diagonal elements

$$P_i^{(\mu)} = \frac{d_\mu}{g}\sum_{R\in\mathcal{G}} D_{ii}^{(\mu)*}(R)O(R) \tag{D.1.42}$$

of this operator behave as projection operators. When acting on an arbitrary function $\psi(\boldsymbol{r})$, they project out the part proportional to the ith basis function of the μth irreducible representation:

$$P_i^{(\mu)}\psi(\boldsymbol{r}) = a_i^{(\mu)}\phi_i^{(\mu)}(\boldsymbol{r})\,. \tag{D.1.43}$$

D.1.8 The Double Group

In addition to the position coordinate $\boldsymbol{r}$, the wavefunction of electrons may also contain the spin variable σ. Writing the wave function as a two-component spinor,

$$\psi(\boldsymbol{r},\sigma) = \begin{pmatrix}\psi_+(\boldsymbol{r})\\ \psi_-(\boldsymbol{r})\end{pmatrix}. \tag{D.1.44}$$

The action of an operator $O(R)$ associated with a coordinate transformation R consists not only in the inclusion of $R^{-1}\boldsymbol{r}$ in the argument of the wavefunction but also in the mixing of the two components:

$$O(R)\psi_\sigma(\boldsymbol{r}) = \sum_{\sigma'} D_{\sigma'\sigma}^{1/2}(R)\psi_{\sigma'}(R^{-1}\boldsymbol{r})\,. \tag{D.1.45}$$

The form of the matrix $D^{1/2}_{\sigma'\sigma}(R)$ valid for arbitrary rotations is given in Appendix F. Here we shall content ourselves with the remark that a rotation through 2π around any axis takes the spinor not into itself but into its negative:

$$O\big(C_{\boldsymbol{n}}(2\pi)\big)\begin{pmatrix}\psi_+(\boldsymbol{r})\\ \psi_-(\boldsymbol{r})\end{pmatrix} = -\begin{pmatrix}\psi_+(\boldsymbol{r})\\ \psi_-(\boldsymbol{r})\end{pmatrix}. \tag{D.1.46}$$

The initial state is recovered only after a rotation through 4π. Likewise, since mirror reflection can be considered as the product of an inversion and a rotation through π, two subsequent mirror reflections take the spinor into its negative. Therefore we introduce a new symmetry operation, denoted by $\overline{E}$, which corresponds to a rotation through 2π around any axis. In the spinor space this is represented by the matrix

$$O\big(\overline{E}\big) = \begin{pmatrix}-1 & 0\\ 0 & -1\end{pmatrix}. \tag{D.1.47}$$

If for each transformation R the operation $\overline{R} = R\overline{E} = \overline{E}R$ is also allowed, then the corresponding operator $O(\overline{R})$ acts on the spinor according to the formula

$$O\big(\overline{R}\big)\psi_\sigma(\boldsymbol{r}) = -\sum_{\sigma'} D^{1/2}_{\sigma'\sigma}(R)\psi_{\sigma'}(R^{-1}\boldsymbol{r})\,. \tag{D.1.48}$$

The transformations R and $\overline{R}$ – as well as the corresponding operators – defined in this way form a group. Since this contains twice as many elements as the group $\mathcal{G}$, this is called the double group[1] $\mathcal{G}'$ of the group $\mathcal{G}$.

The irreducible representations and the character table of this group can be determined using the methods discussed above. Rotations that differ by 2π are associated with the same character in certain representations and with opposite characters in others. They are called single- and double-valued representations, respectively. Double-valued representations become important only when spin–orbit interactions have to be taken into account, too. Otherwise even for electrons it is sufficient to know the single-valued representations.

D.1.9 Continuous Groups

Up to now we have always considered groups that contain a finite number of elements, since crystalline solids are invariant only under discrete translations and discrete rotations. However, since the angular momentum discussed in Appendix F is related to the continuous rotation group, we shall briefly present some concepts specific to continuous groups.

O(n), the orthogonal group of degree n is a group of real $n \times n$ matrices such that

$$\mathsf{A}\mathsf{A}^{\mathrm{T}} = \mathsf{I}\,, \tag{D.1.49}$$

[1] H. Bethe, 1929.

where the transpose of a matrix is defined as $\left(\mathsf{A}^{\mathrm{T}}\right)_{ij} = \mathsf{A}_{ji}$. This is equivalent to defining $\mathrm{O}(n)$ as the group of continuous transformations that conserve the length of vectors in n-dimensional space, since if

$$x_i' = \sum_j A_{ij} x_j \tag{D.1.50}$$

then, on account of the above property of A,

$$\sum_i x_i'^2 = \sum_i x_i^2 \,. \tag{D.1.51}$$

Therefore the group $\mathrm{O}(n)$ contains rotations through an arbitrary angle around axes of arbitrary orientation that pass through a fixed point in n-dimensional space, reflections on the planes that contain the same point, and inversion in the same point.

Owing to the condition imposed on matrix A, $\det(\mathsf{A}) = \pm 1$. By taking only those matrices for which $\det(\mathsf{A}) = 1$, one obtains $\mathrm{SO}(n)$, the special orthogonal group of degree n. This corresponds to pure rotations in n-space.

$\mathrm{U}(n)$, the unitary group of degree n is the group of $n \times n$ complex matrices such that

$$\mathsf{A}\mathsf{A}^\dagger = \mathsf{I}, \tag{D.1.52}$$

where the adjoint matrix is defined as $\left(\mathsf{A}^\dagger\right)_{ij} = \mathsf{A}^*_{ji}$. This is equivalent to defining $\mathrm{U}(n)$ as the group of continuous transformations that conserve the quantity $\sum_i \left|z_i\right|^2$ in n-dimensional complex space, that is

$$\sum_i \left|z_i'\right|^2 = \sum_i \left|z_i\right|^2. \tag{D.1.53}$$

On account of the condition imposed on matrix A, $\det(\mathsf{A}) = \pm 1$ here, too. By taking only those matrices for which $\det(\mathsf{A}) = 1$, one obtains $\mathrm{SU}(n)$, the special unitary group of degree n. In the $n = 2$ case their most general form is

$$\begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix}, \tag{D.1.54}$$

where a and b are complex numbers and $|a|^2 + |b|^2 = 1$, or in another form

$$\begin{pmatrix} \mathrm{e}^{\mathrm{i}\xi} \cos\eta & -\mathrm{e}^{\mathrm{i}\zeta} \sin\eta \\ \mathrm{e}^{-\mathrm{i}\zeta} \sin\eta & \mathrm{e}^{-\mathrm{i}\xi} \cos\eta \end{pmatrix}. \tag{D.1.55}$$

These are special cases of continuous Lie groups. The r-parameter Lie group of coordinate transformations can be given as

$$x_i' = f_i(x_1, \ldots, x_n, a_1, \ldots, a_r) \,. \tag{D.1.56}$$

The number of parameters is $n(n-1)/2$ for $\mathrm{SO}(n)$, n^2 for $\mathrm{U}(n)$, and $n^2 - 1$ for $\mathrm{SU}(n)$ groups. Without loss of generality the parameters a_ν can be

chosen in such a way that their zero value should correspond to the identity transformation. For small values δa_ν of the parameters it is required that up to linear order the transformation could be written as

$$x_i' = \left[1 + \mathrm{i}\sum_{\nu=1}^{r} I_\nu \delta a_\nu\right] x_i \,. \tag{D.1.57}$$

The quantities I_ν ($\nu = 1, \ldots, r$) in the previous expression are the generators of the transformation, since it can be shown that for arbitrary values of the parameters the new coordinates can be expressed in terms of these:

$$x_i' = \exp\left(\mathrm{i}\sum_{\nu=1}^{r} I_\nu a_\nu\right) x_i \,. \tag{D.1.58}$$

As usual, an operator $O(R)$ is associated with the coordinate transformation in such a way that

$$O(R)\psi(\boldsymbol{r}) = \psi(R^{-1}\boldsymbol{r}) \,, \tag{D.1.59}$$

that is for $\psi(\boldsymbol{r}) = \boldsymbol{r}$

$$O(R)\boldsymbol{r} = \left[1 - \mathrm{i}\sum_{\nu=1}^{r} I_\nu \delta a_\nu\right] \boldsymbol{r} \,. \tag{D.1.60}$$

The generators therefore appear as $n \times n$ matrices. For finite values of the parameters,

$$O(R)\boldsymbol{r} = \exp\left(-\mathrm{i}\sum_{\nu=1}^{r} I_\nu a_\nu\right) \boldsymbol{r} \,. \tag{D.1.61}$$

The commutator of two generators is also a generator, and the coefficients $c_{\alpha\beta\gamma}$ present in the commutator

$$\left[I_\alpha, I_\beta\right] = \mathrm{i}\sum_{\gamma=1}^{r} c_{\alpha\beta\gamma} I_\gamma \tag{D.1.62}$$

are the structure constants of the Lie group's Lie algebra.

For rotations in the (x, y) plane, when the transformation is characterized by a single parameter, the angle ϕ of the rotation,

$$\begin{aligned} x' &= x\cos\phi - y\sin\phi \,, \\ y' &= x\sin\phi + y\cos\phi \,, \end{aligned} \tag{D.1.63}$$

a single generator appears:

$$I = \frac{1}{\mathrm{i}}\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right) = \frac{1}{\mathrm{i}}\frac{\partial}{\partial \phi} \,, \tag{D.1.64}$$

since

$$\begin{aligned} \mathrm{e}^{\mathrm{i}\phi I} x &= x\cos\phi - y\sin\phi\,, \\ \mathrm{e}^{\mathrm{i}\phi I} y &= x\sin\phi + y\cos\phi\,. \end{aligned} \tag{D.1.65}$$

Note that I is just the z component of the dimensionless angular momentum operator. Writing the generator as a 2×2 matrix,

$$I = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \tag{D.1.66}$$

and obviously

$$\begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} = \mathrm{e}^{-\mathrm{i}\phi I}. \tag{D.1.67}$$

For the group SO(3) the three generators belong to the infinitesimal rotations around the three axes. Rotation through angle ϕ around the z-axis corresponds to the transformation

$$\begin{aligned} x' &= x\cos\phi - y\sin\phi\,, \\ y' &= x\sin\phi + y\cos\phi\,, \\ z' &= z\,; \end{aligned} \tag{D.1.68}$$

its generator is

$$I_z = \frac{1}{\mathrm{i}}\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right). \tag{D.1.69}$$

The generators of rotations around the x- and y-axes are, likewise,

$$I_x = \frac{1}{\mathrm{i}}\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right), \qquad I_y = \frac{1}{\mathrm{i}}\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right). \tag{D.1.70}$$

It is readily seen that the generators of the rotation group are the same as the components of the angular momentum operator. Their commutator is known,

$$\left[I_\alpha, I_\beta\right] = \mathrm{i}\sum_{\gamma=1}^{3} \epsilon_{\alpha\beta\gamma} I_\gamma\,, \tag{D.1.71}$$

where $\epsilon_{\alpha\beta\gamma}$ is the completely antisymmetric (Levi-Civita) tensor.

Based on (D.1.60), the generators I_α can be represented by 3×3 matrices as

$$I_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -\mathrm{i} \\ 0 & \mathrm{i} & 0 \end{pmatrix}, \quad I_y = \begin{pmatrix} 0 & 0 & \mathrm{i} \\ 0 & 0 & 0 \\ -\mathrm{i} & 0 & 0 \end{pmatrix}, \quad I_z = \begin{pmatrix} 0 & -\mathrm{i} & 0 \\ \mathrm{i} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{D.1.72}$$

The structure constant of the SU(2) group is also the Levi-Civita symbol, and the three generators satisfy the commutation relation

$$\left[I_\alpha, I_\beta\right] = 2\mathrm{i}\sum_{\gamma=1}^{3} \epsilon_{\alpha\beta\gamma} I_\gamma \,, \tag{D.1.73}$$

while for SU(3) the structure constant $c_{\alpha\beta\gamma}$ that relates the eight generators is totally antisymmetric in its three indices. Its nonvanishing components are $c_{123} = 2$, $c_{147} = c_{165} = c_{246} = c_{257} = c_{345} = c_{376} = 1$, and $c_{458} = c_{678} = \sqrt{3}$.

The simplest, fundamental representation of the SU(2) generators are the three 2×2 Pauli matrices,

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{D.1.74}$$

while that of the SU(3) group is given by the eight Gell-Mann matrices:[2]

$$\begin{aligned}
\lambda_1 &= \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_2 = \begin{pmatrix} 0 & -\mathrm{i} & 0 \\ \mathrm{i} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \\
\lambda_4 &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \lambda_5 = \begin{pmatrix} 0 & 0 & -\mathrm{i} \\ 0 & 0 & 0 \\ \mathrm{i} & 0 & 0 \end{pmatrix}, \quad \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \\
\lambda_7 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -\mathrm{i} \\ 0 & \mathrm{i} & 0 \end{pmatrix}, \qquad \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}.
\end{aligned} \tag{D.1.75}$$

D.2 Applications of Group Theory

While in specific calculations it is important to consider the group that corresponds to the true symmetries of the crystal, for simplicity, we shall almost invariably assume cubic symmetry in our examples. Therefore below we shall give the irreducible representations for the cubic O_h group. Then we shall list a number of statements and theorems that arise from the connection between group theory and quantum mechanics, and that are particularly useful in solid-state physics applications.

D.2.1 Irreducible Representations of the Group O_h

The rotations and rotoinversions of the group O_h are listed in Tables 5.1 and 5.4. From the multiplication law for symmetry elements it may be shown that rotations through 90°, 120°, and 180° belong to separate classes – and even

[2] Murray Gell-Mann (1929–) was awarded the Nobel Prize in 1969 "for his contributions and discoveries concerning the classification of elementary particles and their interactions".

among 180° rotations separate classes are formed by rotations around edges and face diagonals. Since the identity element always constitutes a separate class, the 24 rotations can be divided into 5 classes:

1. the identity element E;
2. the 6 fourfold rotations, C_{4m} and C_{4m}^3 $(m = x,\, y,\, z)$;
3. the 3 twofold rotations C_{2m} around the same axes;
4. the 6 twofold rotations C_{2p} $(p = a,\, b,\, c,\, d,\, e,\, f)$ around the face diagonals;
5. the 8 threefold rotations C_{3j} and C_{3j}^2 $(j = a, b, c, d)$ around the space diagonals.

The other 24 elements – which can be obtained from rotations via a multiplication by inversion – form 5 classes analogously: just the elements of the above classes need to be multiplied by the inversion element.

As for finite groups the number of irreducible representations is the same as the number of classes, the group O_h has 10 irreducible representations. Using (D.1.20), the relation for the dimensionality of the irreducible representations, the equation

$$\sum_{i=1}^{10} d_i^2 = 48 \tag{D.2.1}$$

can be satisfied by integers if there are 4 one-, 2 two-, and 4 three-dimensional irreducible representations.

Two conventions are used in the literature for the notation of irreducible representations. The first one is due to L. P. Bouckaert, R. Smoluchowski, and E. P. Wigner, who determined in 1936 the irreducible representations of the groups that are of importance in solid-state physics. A letter denotes the point of the Brillouin zone for which the group in question is the symmetry group. An additional number shows the dimensionality of the representation, and a prime ($'$) if the representation is odd under inversion. In the other notation (called the chemical notation) the letters A, E, and T show the dimensionality of the representation, while the indices g and u show whether the representation is even (in German: *gerade*) or odd (*ungerade*). Using the orthogonality relations (D.1.23) and (D.1.24) of the characters of the irreducible representations, the characters can be determined. The character table of the group O_h is given in Table D.1.

Now consider the irreducible representations of the double group. We have seen that the group O that contains the 24 rotational symmetries of the cube has 5 irreducible representations. However, the corresponding 48-element double group O' has only 8 irreducible representations (i.e., its elements can be divided into 8 classes). Besides the identity element E, the element $\overline{E}$ that corresponds to a rotation through 2π constitutes a group on its own. New classes are formed also by the rotations through $\pi/2 + 2\pi$ and $2\pi/3 + 2\pi$, however rotations through π and $\pi + 2\pi$ belong to the same class. Of the 8 representations 5 are single-valued – these are the same as the representations

Table D.1. The character table of the irreducible representations of the group O_h

		E	$3C_{4m}^2$	$6C_{4m}$	$6C_{2p}$	$8C_{3j}$	I	$3\sigma_m$	$6S_{4m}$	$6\sigma_p$	$8S_{3j}$
Γ_1	A_{1g}	1	1	1	1	1	1	1	1	1	1
Γ_2	A_{2g}	1	1	−1	−1	1	1	1	−1	−1	1
Γ_{12}	E_g	2	2	0	0	−1	2	2	0	0	−1
Γ'_{15}	T_{1g}	3	−1	1	−1	0	3	−1	1	−1	0
Γ'_{25}	T_{2g}	3	−1	−1	1	0	3	−1	−1	1	0
Γ'_1	A_{1u}	1	1	1	1	1	−1	−1	−1	−1	−1
Γ'_2	A_{2u}	1	1	−1	−1	1	−1	−1	1	1	−1
Γ'_{12}	E_u	2	2	0	0	−1	−2	−2	0	0	1
Γ_{15}	T_{1u}	3	−1	1	−1	0	−3	1	−1	1	0
Γ_{25}	T_{2u}	3	−1	−1	1	0	−3	1	1	−1	0

of the group O –, and 3 are double-valued. The character table of the double group is given in Table D.2.

Table D.2. Character table of the irreducible representations for the double group O' of the point group O

		E	$\overline{E}$	$3C_{4m}^2$ $+3\overline{C}_{4m}^2$	$6C_{4m}$	$6\overline{C}_{4m}$	$6C_{2p}$ $+6\overline{C}_{2p}$	$8C_{3j}$	$8\overline{C}_{3j}$
Γ_1	A_1	1	1	1	1	1	1	1	1
Γ_2	A_2	1	1	1	−1	−1	−1	1	1
Γ_3	E	2	2	2	0	0	0	−1	−1
Γ_4	T_1	3	3	−1	1	1	−1	0	0
Γ_5	T_2	3	3	−1	−1	−1	1	0	0
Γ_6	$\overline{E}_1$	2	−2	0	$\sqrt{2}$	$-\sqrt{2}$	0	1	−1
Γ_7	$\overline{E}_2$	2	−2	0	$-\sqrt{2}$	$\sqrt{2}$	0	1	−1
Γ_8	$\overline{U}$	4	−4	0	0	0	0	−1	1

D.2.2 Group Theory and Quantum Mechanics

The Hamiltonian of a quantum mechanical system may be invariant under certain coordinate transformations R, that is for each point $\boldsymbol{r}$ in space the relation

$$\mathcal{H}(\boldsymbol{r}') = \mathcal{H}(\boldsymbol{r})\,, \qquad \text{if} \quad \boldsymbol{r}' = R\boldsymbol{r} \tag{D.2.2}$$

may be satisfied. The coordinate transformations R that meet this requirement form a group $\mathcal{G}$. This is called the symmetry group of the Hamiltonian. Invariance under the elements of the group $\mathcal{G}$ also means that the Hamiltonian commutes with the operator $O(R)$ associated with the coordinate transformation R, since

$$O(R)\left[\mathcal{H}(\boldsymbol{r})\psi(\boldsymbol{r})\right] = \mathcal{H}(R^{-1}\boldsymbol{r})\psi(R^{-1}\boldsymbol{r}) = \mathcal{H}(\boldsymbol{r})O(R)\psi(\boldsymbol{r})\,, \tag{D.2.3}$$

and hence

$$O(R)\mathcal{H}(\boldsymbol{r}) = \mathcal{H}(\boldsymbol{r})O(R)\,. \tag{D.2.4}$$

This also implies that if $\psi(\boldsymbol{r})$ is an eigenfunction of the Hamiltonian with energy ε then $O(R)\psi(\boldsymbol{r})$ is also an eigenfunction, with the same energy. This is Wigner's theorem, which was discussed in Chapter 6.

The irreducible representations of the symmetry group of the Hamiltonian and the basis functions of the representations play a privileged role in the solution of the quantum mechanical eigenvalue problem. Below we shall list without proof a couple of theorems which are, to a certain extent, different formulations of the same statement.

If the Hamiltonian $\mathcal{H}$ is invariant under the transformations of a symmetry group $\mathcal{G}$ then the eigenfunctions that belong to the same energy form the basis of a representation of the group $\mathcal{G}$. For an n-fold degenerate energy level there exist n linearly independent eigenfunctions $\psi_i(\boldsymbol{r})$; the representation is provided by the matrices $\mathsf{D}(R)$ that appear in the equation

$$O(R)\psi_i(\boldsymbol{r}) = \sum_j D_{ji}(R)\psi_j(\boldsymbol{r}) \tag{D.2.5}$$

specifying their transformation properties. This representation is usually reducible. If, however, the group $\mathcal{G}$ contains every possible symmetry of the Hamiltonian then – barring accidental degeneracies – for each energy level the corresponding eigenfunctions transform according to an irreducible representation of the group $\mathcal{G}$.

In quantum mechanical calculations one often needs to determine the matrix elements

$$M_{ij} = \left(\phi_i^{(\mu)}, Q\phi_j^{(\nu)}\right) \tag{D.2.6}$$

of an operator Q, where $\phi_i^{(\mu)}(\boldsymbol{r})$ and $\phi_j^{(\nu)}(\boldsymbol{r})$ are the basis functions of the irreducible representations $D^{(\mu)}$ and $D^{(\nu)}$. If the symmetry of the operator Q is lower than that of the Hamiltonian then the operators $O(R)$ that belong to the elements of the group $\mathcal{G}$ take the operator Q into

$$Q_i \equiv O(R)QO^{-1}(R)\,. \tag{D.2.7}$$

The operators Q_i obtained this way are transformed into each other according to a representation D that is not the identity representation:

$$O(R)Q_iO^{-1}(R) = \sum_j D_{ij}(R)Q_j\,. \tag{D.2.8}$$

It can be shown that the matrix element in (D.2.6) vanishes unless the representation $D^{(\mu)}$ appears in the reduction of the direct product of D and $D^{(\nu)}$. In other words: the matrix element vanishes unless the identity representation appears in the reduction of the direct product of ${D^{(\mu)}}^*$, D, and $D^{(\nu)}$. Using these *selection rules* it is possible to determine whether a quantum mechanical transition occurs.

If the operator Q possesses the full symmetry of the Hamiltonian then the matrix element is finite provided the identity representation appears in the reduction of the direct product ${D^{(\mu)}}^* \otimes D^{(\nu)}$. As a special case consider the matrix elements of the Hamiltonian between the basis functions of irreducible representations. If $D^{(\mu)}(R)$ and $D^{(\nu)}(R)$ are two irreducible unitary representations of the symmetry group of the Hamiltonian, and the associated basis functions are $\phi_1^{(\mu)}(\boldsymbol{r})$, $\phi_2^{(\mu)}(\boldsymbol{r})$, ... and $\phi_1^{(\nu)}(\boldsymbol{r})$, $\phi_2^{(\nu)}(\boldsymbol{r})$, ... then

$$\left(\phi_i^{(\mu)}, \mathcal{H}\phi_j^{(\nu)}\right) = \varepsilon^{(\mu)}\delta_{\mu\nu}\delta_{ij}\,, \tag{D.2.9}$$

that is, the Hamiltonian has only diagonal matrix elements in this basis, and the matrix element is the same for each basis function of the μth irreducible representation. This means that if the Hamiltonian $\mathcal{H}$ is invariant under the transformations of a group $\mathcal{G}$ then the eigenfunctions that transform according to the same irreducible representation are of the same energy.

As a consequence of this, consider a Hamiltonian that may be written in the form $\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_1$, where $\mathcal{H}_1$ can be taken as a perturbation. Assume further that the symmetry group of $\mathcal{H}_0$ is $\mathcal{G}_0$, while $\mathcal{H}_1$ is invariant only under the transformations of a subgroup $\mathcal{G}$ of $\mathcal{G}_0$. Then under the perturbation $\mathcal{H}_1$ any energy level ε_0 of $\mathcal{H}_0$ will split so that when the (usually reducible) representation of the group $\mathcal{G}$ over the eigenfunctions of the level ε_0 is reduced only those states will necessarily be of the same energy that belong to the same irreducible representation. The energy level ε_0 will be split into at most as many levels of different energy as the number of irreducible representations that appear in the reduction (and that are compatible with the representation of the level ε_0). If the representation of the group $\mathcal{G}$ over the eigenfunctions of the level ε_0 is irreducible then the perturbation only shifts but does not split the level.

References

1. J. E. Cornwell, *Group Theory in Physics*, Vol. 1, Fourth Printing, Academic Press, London (1989).
2. A. P. Cracknell, *Group Theory in Solid State Physics*, Taylor & Francis Ltd., London (1975).

3. M. Hamermesh, *Group Theory and Its Applications to Physical Problems*, Addison-Wesley/Pergamon, New York (1962).
4. R. L. Liboff, *Primer for Point and Space Groups*, Undergraduate Texts in Contemporary Physics, Springer-Verlag, Berlin (2004).
5. M. Tinkham, *Group Theory and Quantum Mechanics*, McGraw-Hill, New York (1964).
6. E. P. Wigner, *Group Theory and Its Applications to the Quantum Mechanics of Atomic Spectra*, Academic Press, New York (1959).

E

Scattering of Particles by Solids

A frequently used method for the investigation of solids is the scattering of particles (electrons, photons, neutrons, etc.) by the solid. In an ideal setup well-collimated monoenergetic beams are used in which the energy and the wave vector of the incident particles are known. By measuring the energy of the particles scattered in various directions (i.e., the energy distribution), the changes in energy and momentum permit us to determine the dispersion relation of the excitations created or absorbed in the sample, and through it to study the internal dynamics of the system.

By assuming potential scattering and magnetic interactions, we shall derive the general formula for the scattering cross section expressed in terms of some correlation function of the system. For potential scattering this will be the density–density correlation function, while for magnetic interactions it will be the correlation function for magnetic density.

E.1 The Scattering Cross Section

For simplicity, assume that the incoming and scattered beams can be both approximated by plane waves, of wave vectors $\boldsymbol{k}$ and $\boldsymbol{k}'$, respectively. The properly normalized wavefunctions are

$$|\boldsymbol{k}\rangle = \frac{1}{\sqrt{V}} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\,, \qquad |\boldsymbol{k}'\rangle = \frac{1}{\sqrt{V}} \mathrm{e}^{\mathrm{i}\boldsymbol{k}'\cdot\boldsymbol{r}}\,, \tag{E.1.1}$$

where V is the total volume, not only that of the sample. Assuming that the interaction with the scattered particle gives rise to the creation or absorption of an elementary excitation in the sample, and using the consequences of discrete translational symmetries derived in Chapter 6, the wave vector of the elementary excitation created in the scattering process between states $|\boldsymbol{k}\rangle$ and $|\boldsymbol{k}'\rangle$ is

$$\boldsymbol{q} = \boldsymbol{k} - \boldsymbol{k}' + \boldsymbol{G}\,. \tag{E.1.2}$$

In the absorption process an elementary excitation of wave vector $-\boldsymbol{q}$ is absorbed. $\boldsymbol{G}$ is a vector of the reciprocal lattice.

Simultaneously, the crystal makes a transition from the initial state $|i\rangle$ of energy E_i to a final state $|f\rangle$ of energy E_f. The energy of the elementary excitation is the difference of these two. On account of energy conservation, the change in the energy of the sample can be expressed with the change in the particle energy. For the scattering of particles of mass m_n with a quadratic dispersion relation, the energy transferred to the sample in the scattering process is

$$\varepsilon = \frac{\hbar^2 k^2}{2m_\mathrm{n}} - \frac{\hbar^2 k'^2}{2m_\mathrm{n}} = E_\mathrm{f} - E_\mathrm{i}\,. \tag{E.1.3}$$

If a single elementary excitation is created or annihilated in the scattering process, then an inelastic peak appears in the distribution of the scattered particles. Its direction is determined by the momentum of the excitation, while its energy is specified by the energy transfer in the process. If the changes in the momentum and energy of the scattered particle can be measured simultaneously, then the dispersion relation for the elementary excitation created or annihilated in the process can be determined.

The situation is often not so simple in scattering experiments, therefore evaluating the scattering cross section, which determines the strength of the scattering process is of the utmost importance.

Suppose that scattering takes the sample from a well-defined initial state $|i\rangle$ to an equally well defined final state $|f\rangle$. The joint state of the sample and the particle is denoted by $|i\rangle|\boldsymbol{k}\rangle$ before scattering, and by $|f\rangle|\boldsymbol{k}'\rangle$ after it. Assume, furthermore, that the Hamiltonian $\mathcal{H}_\mathrm{int}$ of the interaction between the particle and the crystal is known.

Owing to the weakness of the interaction, the Born approximation suffices to determine the scattering cross section in the majority of cases. According to quantum mechanics, the transition probability between the initial state and a specified final state is

$$W_{\mathrm{i},\boldsymbol{k}\to\mathrm{f},\boldsymbol{k}'} = \frac{2\pi}{\hbar}\left|\langle\boldsymbol{k}'|\langle f|\mathcal{H}_\mathrm{int}|i\rangle|\boldsymbol{k}\rangle\right|^2 \delta(\varepsilon - E_\mathrm{f} + E_\mathrm{i})\,, \tag{E.1.4}$$

where the factor $\delta(\varepsilon - E_\mathrm{f} + E_\mathrm{i})$ ensures the conservation of energy.

We shall follow the convention that separates the factor $1/V^{1/2}$ arising from the normalization of the incident and scattered wavefunctions, i.e., uses the notations $|\boldsymbol{k}\rangle = \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}$ and $|\boldsymbol{k}'\rangle = \mathrm{e}^{\mathrm{i}\boldsymbol{k}'\cdot\boldsymbol{r}}$ for the expressions without this normalization factor. Therefore

$$W_{\mathrm{i},\boldsymbol{k}\to\mathrm{f},\boldsymbol{k}'} = \frac{2\pi}{\hbar}\frac{1}{V^2}\left|\langle\boldsymbol{k}'|\langle f|\mathcal{H}_\mathrm{int}|i\rangle|\boldsymbol{k}\rangle\right|^2 \delta(\varepsilon - E_\mathrm{f} + E_\mathrm{i})\,. \tag{E.1.5}$$

The energy and propagation direction of the scattered particle can be measured only with a certain precision. Therefore we have to sum the contributions of those states with wave vector $\boldsymbol{k}'$ whose energy is in the interval

$$\mathrm{d}\varepsilon = \frac{\mathrm{d}\varepsilon}{\mathrm{d}k'}\,\mathrm{d}k' = \frac{\hbar^2 k'}{m_\mathrm{n}}\,\mathrm{d}k' \tag{E.1.6}$$

and whose direction is in the element of solid angle $\mathrm{d}\Omega = \sin\theta\,\mathrm{d}\theta\,\mathrm{d}\varphi$. The number of such states is given by

$$\rho(\varepsilon)\mathrm{d}\varepsilon\,\mathrm{d}\Omega = \frac{V}{(2\pi)^3}k'^2\,\mathrm{d}k'\,\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\varphi = V\frac{k' m_\mathrm{n}}{(2\pi)^3\hbar^2}\,\mathrm{d}\varepsilon\,\mathrm{d}\Omega\,. \tag{E.1.7}$$

The intensity of the beam scattered into the selected solid angle element and energy range is obtained by multiplying the transition probability by this density of final states and the number $N_{\boldsymbol{k}}$ of incoming particles in state $\boldsymbol{k}$:

$$\Delta I_{\boldsymbol{k}} = N_{\boldsymbol{k}} V \frac{k' m_\mathrm{n}}{(2\pi)^3\hbar^2}\frac{2\pi}{\hbar}\frac{1}{V^2}\left|\langle \boldsymbol{k}'|\langle f|\mathcal{H}_\mathrm{int}|i\rangle|\boldsymbol{k}\rangle\right|^2 \delta(\varepsilon - E_\mathrm{f} + E_\mathrm{i})\,\mathrm{d}\varepsilon\,\mathrm{d}\Omega\,. \tag{E.1.8}$$

The doubly differential cross section is then obtained through division by the incoming flux Φ. This flux – that is, the number of particles passing through a unit surface in unit time – is just the product of the density of particles in state $\boldsymbol{k}$ and their velocity:

$$\Phi = \frac{N_{\boldsymbol{k}}}{V}\frac{\hbar k}{m_\mathrm{n}}\,. \tag{E.1.9}$$

Thus the volume factors eventually cancel out, and the cross section is

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}\Omega\,\mathrm{d}\varepsilon} = \frac{k'}{k}\left(\frac{m_\mathrm{n}}{2\pi\hbar^2}\right)^2\left|\langle \boldsymbol{k}'|\langle f|\mathcal{H}_\mathrm{int}|i\rangle|\boldsymbol{k}\rangle\right|^2 \delta(\varepsilon - E_\mathrm{f} + E_\mathrm{i})\,. \tag{E.1.10}$$

Very similar expressions are derived for photon scattering, only the prefactors of the matrix element are slightly different: owing to the different dispersion relation, neither the density of final states nor the incident flux are the same as before. Using the angular frequency instead of the energy, the density of final states, derived from

$$\rho(\omega)\mathrm{d}\omega\,\mathrm{d}\Omega = \frac{V}{(2\pi)^3}k'^2\,\mathrm{d}k'\,\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\varphi\,, \tag{E.1.11}$$

is now

$$\rho(\omega) = \frac{V}{(2\pi)^3}k'^2\frac{\mathrm{d}k'}{\mathrm{d}\omega} = \frac{V}{(2\pi)^3}k'^2\frac{n}{c}\,, \tag{E.1.12}$$

where n is the refractive index of the sample. Writing the incoming photon flux as

$$\Phi = \frac{v}{V} = \frac{c}{nV}\,, \tag{E.1.13}$$

and expressing it in terms of the change in frequency rather than the change in energy, the cross section is now

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}\Omega\,\mathrm{d}\omega} = \left(\frac{nk'}{2\pi\hbar c}\right)^2\left|\langle \boldsymbol{k}'|\langle f|\mathcal{H}_\mathrm{int}|i\rangle|\boldsymbol{k}\rangle\right|^2 \delta(\omega - (E_\mathrm{f} - E_\mathrm{i})/\hbar)\,. \tag{E.1.14}$$

Below we shall derive the formulas for neutron scattering, however, our results will apply – up to a multiplicative factor – to light scattering as well.

The expression studied above is for fixed initial and final states of the sample. However, neither is known from measurements, which provide data only about the incident and scattered particles. Assuming that the sample is initially in thermal equilibrium, the probability of the initial state $|i\rangle$ is

$$p_i = \frac{1}{Z} \mathrm{e}^{-\beta E_i} \,. \tag{E.1.15}$$

This probability has to be used as the weighting factor of the initial states in the averaging procedure. Finally, the final states $|f\rangle$ for which the transition matrix element is finite must be summed over. The doubly differential cross section is then

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}\Omega\,\mathrm{d}\varepsilon} = \frac{k'}{k} \left(\frac{m_\mathrm{n}}{2\pi\hbar^2}\right)^2 \sum_{if} p_\mathrm{i} \left|\langle \boldsymbol{k}'|\langle f|\mathcal{H}_\mathrm{int}|i\rangle|\boldsymbol{k}\rangle\right|^2 \delta(\varepsilon - E_\mathrm{f} + E_\mathrm{i}) \,. \tag{E.1.16}$$

E.2 The Van Hove Formula for Cross Section

As it was pointed out by L. Van Hove in 1954, the above form of the scattering cross section can be expressed with some correlation function of the scattering centers, the exact form of which depends on the type of the interaction between the scattered particles and the sample.

E.2.1 Potential Scattering

The previous general expression will be reformulated by assuming that the interaction between the scattered particle and the electrons or ions of the crystal can be written as

$$\mathcal{H}_\mathrm{int} = \sum_m U(\boldsymbol{r} - \boldsymbol{r}_m) \,, \tag{E.2.1}$$

where $\boldsymbol{r}_m$ stands for the position coordinates of the mth scattering center in the crystal. Substituting this expression into the transition matrix element, and choosing the wavefunction of the scattered particles as a plane wave,

$$\begin{aligned} \langle \boldsymbol{k}|\langle i|\mathcal{H}_\mathrm{int}|f\rangle|\boldsymbol{k}'\rangle &= \int \mathrm{d}\boldsymbol{r} \left\langle i\right| \mathrm{e}^{-\mathrm{i}(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{r}} \sum_m U(\boldsymbol{r} - \boldsymbol{r}_m) \left|f\right\rangle \\ &= U_{\boldsymbol{K}} \left\langle i\right| \sum_m \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}_m} \left|f\right\rangle \end{aligned} \tag{E.2.2}$$

is obtained, where $\boldsymbol{K} = \boldsymbol{k} - \boldsymbol{k}'$ and $\hbar\boldsymbol{K}$ is the momentum transferred in the scattering process, while

$$U_{\boldsymbol{K}} = \int \mathrm{d}\boldsymbol{r}\, \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}} U(\boldsymbol{r}) \tag{E.2.3}$$

is the Fourier transform of the interaction at the wave number that corresponds to the above momentum transfer.

Substituting this into the cross section formula (E.1.16),

$$\begin{aligned}\frac{\mathrm{d}^2\sigma}{\mathrm{d}\Omega\,\mathrm{d}\varepsilon} = \frac{k'}{k}\left(\frac{m_\mathrm{n}}{2\pi\hbar^2}\right)^2 |U_{\boldsymbol{K}}|^2 \sum_{i,f}\sum_{m,n} p_\mathrm{i} \left\langle i \left| \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}_m} \right| f \right\rangle \\ \times \left\langle f \left| \mathrm{e}^{\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}_n} \right| i \right\rangle \delta(\varepsilon - E_\mathrm{f} + E_\mathrm{i})\,.\end{aligned} \tag{E.2.4}$$

Using the notation

$$A_{\boldsymbol{K}} = \frac{k'}{k}\left(\frac{m_\mathrm{n}}{2\pi\hbar^2}\right)^2 |U_{\boldsymbol{K}}|^2\,, \tag{E.2.5}$$

the cross section is written as

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}\Omega\,\mathrm{d}\varepsilon} = A_{\boldsymbol{K}} \sum_{i,f}\sum_{m,n} p_i \left\langle i \left| \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}_m} \right| f \right\rangle \left\langle f \left| \mathrm{e}^{\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}_n} \right| i \right\rangle \delta(\varepsilon - E_\mathrm{f} + E_\mathrm{i})\,. \tag{E.2.6}$$

Making use of the integral representation (C.3.1-d) of the Dirac delta function,

$$\delta(\varepsilon - E_\mathrm{f} + E_\mathrm{i}) = \frac{1}{2\pi\hbar}\int\limits_{-\infty}^{\infty} \mathrm{d}t\, \mathrm{e}^{\mathrm{i}(\varepsilon - E_\mathrm{f} + E_\mathrm{i})t/\hbar}\,. \tag{E.2.7}$$

A suitable rearrangement of the terms in the exponent leads to

$$\begin{aligned}\frac{\mathrm{d}^2\sigma}{\mathrm{d}\Omega\,\mathrm{d}\varepsilon} = A_{\boldsymbol{K}}\frac{1}{2\pi\hbar}\int\limits_{-\infty}^{\infty} \mathrm{d}t \sum_{i,f}\sum_{m,n} p_i \left\langle i \left| \mathrm{e}^{\mathrm{i}E_i t/\hbar}\, \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}_m}\, \mathrm{e}^{-\mathrm{i}E_\mathrm{f} t/\hbar} \right| f \right\rangle \\ \times \left\langle f \left| \mathrm{e}^{\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}_n} \right| i \right\rangle \mathrm{e}^{\mathrm{i}\varepsilon t/\hbar}\,.\end{aligned} \tag{E.2.8}$$

Expressing the time-dependent factors of the first matrix element in terms of the Hamiltonian $\mathcal{H}_0$ of the sample, and using the usual form of the time-dependent operators,

$$\begin{aligned}\frac{\mathrm{d}^2\sigma}{\mathrm{d}\Omega\,\mathrm{d}\varepsilon} &= A_{\boldsymbol{K}}\frac{1}{2\pi\hbar}\int\limits_{-\infty}^{\infty} \mathrm{d}t\, \mathrm{e}^{\mathrm{i}\varepsilon t/\hbar} \sum_{i,f}\sum_{m,n} p_i \left\langle i \left| \mathrm{e}^{\mathrm{i}\mathcal{H}_0 t/\hbar}\, \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}_m}\, \mathrm{e}^{-\mathrm{i}\mathcal{H}_0 t/\hbar} \right| f \right\rangle \\ &\quad \times \left\langle f \left| \mathrm{e}^{\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}_n} \right| i \right\rangle \\ &= A_{\boldsymbol{K}}\frac{1}{2\pi\hbar}\int\limits_{-\infty}^{\infty} \mathrm{d}t\, \mathrm{e}^{\mathrm{i}\varepsilon t/\hbar} \sum_{i,f}\sum_{m,n} p_i \left\langle i \left| \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}_m(t)} \right| f \right\rangle \left\langle f \left| \mathrm{e}^{\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}_n} \right| i \right\rangle .\end{aligned} \tag{E.2.9}$$

Making use of the completeness relation for the states, the sum over final states can be evaluated. Using the notation of thermal averaging for denoting summation over the initial states multiplied by the weight factors p_i,

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}\Omega\,\mathrm{d}\varepsilon} = A_{\boldsymbol{K}}\frac{1}{2\pi\hbar}\int\limits_{-\infty}^{\infty}\mathrm{d}t\,\mathrm{e}^{\mathrm{i}\varepsilon t/\hbar}\sum_{m,n}\left\langle \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}_m(t)}\mathrm{e}^{\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}_n}\right\rangle . \tag{E.2.10}$$

Using the identity

$$\left\langle \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}_m(t)}\,\mathrm{e}^{\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}_n}\right\rangle = \left\langle \int \mathrm{d}\boldsymbol{r}\,\mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}}\delta(\boldsymbol{r}-\boldsymbol{r}_m(t))\int \mathrm{d}\boldsymbol{r}'\,\mathrm{e}^{\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}'}\delta(\boldsymbol{r}'-\boldsymbol{r}_n)\right\rangle \tag{E.2.11}$$

and the density

$$\rho(\boldsymbol{r},t) = \sum_m \delta(\boldsymbol{r}-\boldsymbol{r}_m(t)) \tag{E.2.12}$$

of the scattering system, the thermal average in the cross section can be rewritten as

$$\sum_{m,n}\left\langle \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}_m(t)}\,\mathrm{e}^{\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}_n}\right\rangle = \iint \mathrm{d}\boldsymbol{r}\,\mathrm{d}\boldsymbol{r}'\,\mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot(\boldsymbol{r}-\boldsymbol{r}')}\left\langle\rho(\boldsymbol{r},t)\rho(\boldsymbol{r}',0)\right\rangle . \tag{E.2.13}$$

This formula contains the density–density correlation function of scattering centers.

In line with the normalization introduced in Chapter 2,

$$\Gamma(\boldsymbol{r},\boldsymbol{r}',t) = \frac{V}{N}\left\langle\rho(\boldsymbol{r},t)\rho(\boldsymbol{r}',0)\right\rangle , \tag{E.2.14}$$

where N is the number of scattering centers. For spatially homogeneous systems $\Gamma(\boldsymbol{r},\boldsymbol{r}',t)$ is a function of $\boldsymbol{r}-\boldsymbol{r}'$ alone, therefore it is customary to use the notation

$$\Gamma(\boldsymbol{r},t) = \frac{V}{N}\left\langle\rho(\boldsymbol{r}+\boldsymbol{r}',t)\rho(\boldsymbol{r}',0)\right\rangle , \tag{E.2.15}$$

which gives

$$\sum_{m,n}\left\langle \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}_m(t)}\,\mathrm{e}^{\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}_n}\right\rangle = \iint \mathrm{d}\boldsymbol{r}\,\mathrm{d}\boldsymbol{r}'\,\mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}}\left\langle\rho(\boldsymbol{r}+\boldsymbol{r}',t)\rho(\boldsymbol{r}',0)\right\rangle . \tag{E.2.16}$$

Since the integrand is in fact independent of $\boldsymbol{r}'$, integration with respect to $\boldsymbol{r}'$ gives the volume of the sample, so

$$\sum_{m,n}\left\langle \mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}_m(t)}\,\mathrm{e}^{\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}_n}\right\rangle = N\int \mathrm{d}\boldsymbol{r}\,\mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}}\Gamma(\boldsymbol{r},t) . \tag{E.2.17}$$

Introducing the spatial and temporal Fourier transform of the density–density correlation function $\Gamma(\boldsymbol{r},t)$ by

$$S(\boldsymbol{K},\varepsilon/\hbar) = \int\limits_{-\infty}^{\infty}\mathrm{d}t\,\mathrm{e}^{\mathrm{i}\varepsilon t/\hbar}\int \mathrm{d}\boldsymbol{r}\,\mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}}\Gamma(\boldsymbol{r},t) , \tag{E.2.18}$$

the cross section is written as

$$\boxed{\frac{\mathrm{d}^2\sigma}{\mathrm{d}\Omega\,\mathrm{d}\varepsilon} = \frac{N}{2\pi\hbar} A_{\boldsymbol{K}} S(\boldsymbol{K}, \varepsilon/\hbar)\,.} \tag{E.2.19}$$

Looking back at (E.2.5), the previous expression is factorized into three terms: $\frac{k'}{k}\left(\frac{m_\mathrm{n}}{2\pi\hbar^2}\right)^2$, which is specific to the scattered particle alone, $|U_{\boldsymbol{K}}|^2$ which describes the interaction, and $S(\boldsymbol{K}, \varepsilon/\hbar)$, the *dynamical structure factor*, which depends only on the internal dynamics of the scattering system. The above expression for the cross section is called the *Van Hove formula*. It is very generally valid, since only the applicability of the Born approximation has been assumed. Whether the correlation function of nuclear positions, electronic charge density, or spin density appears in the cross section is determined by the dominant interaction between the incoming particles and the sample.

In nonmagnetic scattering neutrons interact with nuclei at positions $\boldsymbol{r}_m$. Using the Fermi pseudopotential

$$U(\boldsymbol{r}) = \frac{2\pi\hbar^2}{m_\mathrm{n}} a_m \delta(\boldsymbol{r} - \boldsymbol{r}_m) \tag{E.2.20}$$

for characterizing the interaction, the point-like character of the scattering center makes the Fourier transform independent of $\boldsymbol{K}$:

$$U_{\boldsymbol{K}} = \frac{2\pi\hbar^2}{m_\mathrm{n}} a_m\,. \tag{E.2.21}$$

For identical atoms the factor $A_{\boldsymbol{K}}$ takes the $\boldsymbol{K}$-independent form $(k/k')|a_m|^2$, and the Fourier transform of the correlation function for the density of the nuclei appears in the cross section. From this formula information can be obtained about nuclear positions, about the equilibrium atomic positions in the crystal, and, through the motion of the nuclei, about phonons. This is discussed in more detail in Chapter 13.

Expression (E.2.10) for the cross section is valid also for the scattering of electrons by crystals. Through electronic position coordinates the spatial distribution of electrons bound to atoms appear in this case. If the electronic wavefunction is $\phi_\mathrm{i}(\boldsymbol{r})$ in the initial and $\phi_\mathrm{f}(\boldsymbol{r})$ in the final state, the transition matrix element is

$$\langle \boldsymbol{k}, i\,|\mathcal{H}_\mathrm{int}|\,f, \boldsymbol{k}'\rangle = \int \mathrm{d}\boldsymbol{r}_\mathrm{n}\,\mathrm{d}\boldsymbol{r}\phi_\mathrm{i}^*(\boldsymbol{r})\phi_\mathrm{f}(\boldsymbol{r})\mathrm{e}^{-\mathrm{i}(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{r}_\mathrm{n}} U(\boldsymbol{r} - \boldsymbol{r}_\mathrm{n})\,. \tag{E.2.22}$$

When the relative coordinate is used, then, besides the Fourier transform of the potential, an additional factor

$$F(\boldsymbol{K}) = \int \mathrm{d}\boldsymbol{r}\phi_\mathrm{i}^*(\boldsymbol{r})\phi_\mathrm{f}(\boldsymbol{r})\mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}} \tag{E.2.23}$$

appears. If there is no change in the state of the electrons then $F(\boldsymbol{K})$ is the Fourier transform of the electron density inside the atom. If the electron is around an atom at $\boldsymbol{r}_m$, that is, if the wavefunctions are in fact functions of $\boldsymbol{r} - \boldsymbol{r}_m$, then

$$F(\boldsymbol{K})\mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}_m} \tag{E.2.24}$$

appears instead of $F(\boldsymbol{K})$. The structure constant that is left behind after the separation of $|F(\boldsymbol{K})|$ in the transition probability is now determined only by the structure and motion of the centers of mass of atomic electron clouds. Similarly to potential scattering, this leads to Bragg peaks and phonon peaks. From $|F(\boldsymbol{G})|$ – the expression specifying the dependence of the Bragg peak intensities on the reciprocal-lattice vectors $\boldsymbol{G}$ – the spatial distribution of the electrons around the atom can be inferred.

As it was mentioned in Section 13.3, if the scattering amplitude is not the same for each scattering center then, in addition to coherent scattering, an additional incoherent contribution appears. This contribution accounts for interference that occurs not between rays scattered by different atoms but between rays scattered by the same atom at different times. Consequently the cross section of incoherent scattering is proportional to the Fourier transform of the self-correlation function of the atoms.

E.2.2 Magnetic Scattering

Neutrons can be scattered not only by nuclei but, on account of their magnetic moment, by electrons, too. The magnetic field of a neutron of magnetic moment $\boldsymbol{\mu}_\mathrm{n}$ at $\boldsymbol{r}_\mathrm{n}$ can be given by the vector potential $\boldsymbol{A}(\boldsymbol{r})$. Either of the three equivalent forms

$$\begin{aligned}\boldsymbol{A}(\boldsymbol{r}) &= \frac{\mu_0}{4\pi}\frac{\boldsymbol{\mu}_\mathrm{n}\times(\boldsymbol{r}-\boldsymbol{r}_\mathrm{n})}{|\boldsymbol{r}-\boldsymbol{r}_\mathrm{n}|^3}\\ &= -\frac{\mu_0}{4\pi}\boldsymbol{\mu}_\mathrm{n}\times\left(\operatorname{grad}\frac{1}{|\boldsymbol{r}-\boldsymbol{r}_\mathrm{n}|}\right)\\ &= \frac{\mu_0}{4\pi}\operatorname{curl}\frac{\boldsymbol{\mu}_\mathrm{n}}{|\boldsymbol{r}-\boldsymbol{r}_\mathrm{n}|}\end{aligned} \tag{E.2.25}$$

can be used. The derived magnetic induction is

$$\boldsymbol{B}(\boldsymbol{r}) = \operatorname{curl}\boldsymbol{A}(\boldsymbol{r})\,. \tag{E.2.26}$$

If the magnetic moment of the jth electron, located at $\boldsymbol{r}_j$, is $\boldsymbol{\mu}_j$, then the interaction of such moments with the magnetic field can be given by the Hamiltonian

$$\mathcal{H}_\mathrm{int} = -\sum_j \boldsymbol{\mu}_j\cdot\boldsymbol{B}(\boldsymbol{r}_j)\,. \tag{E.2.27}$$

Considering a single term of this sum, and introducing the relative coordinate $\boldsymbol{r} = \boldsymbol{r}_j - \boldsymbol{r}_\mathrm{n}$, some simple algebra leads to the following formula for the interaction with the magnetic moment of the jth electron:

$$\begin{aligned}\mathcal{H}_{\rm int} &= -\frac{\mu_0}{4\pi}\boldsymbol{\mu}_j\cdot\left(\operatorname{curl}\operatorname{curl}\frac{\boldsymbol{\mu}_{\rm n}}{|\boldsymbol{r}|}\right)\\ &= -\frac{\mu_0}{4\pi}\left[\left(\boldsymbol{\mu}_j\cdot\boldsymbol{\nabla}\right)\left(\boldsymbol{\mu}_{\rm n}\cdot\boldsymbol{\nabla}\right)\right]\frac{1}{r}+\frac{\mu_0}{4\pi}\left(\boldsymbol{\mu}_j\cdot\boldsymbol{\mu}_{\rm n}\right)\boldsymbol{\nabla}^2\frac{1}{r}\,.\end{aligned} \tag{E.2.28}$$

For $r\neq 0$ the first term yields the well-known dipole–dipole interaction:

$$\mathcal{H}_{\rm dipole}=\frac{\mu_0}{4\pi}\left[\frac{\boldsymbol{\mu}_j\cdot\boldsymbol{\mu}_{\rm n}}{r^3}-\frac{3(\boldsymbol{\mu}_j\cdot\boldsymbol{r})(\boldsymbol{\mu}_{\rm n}\cdot\boldsymbol{r})}{r^5}\right]. \tag{E.2.29}$$

An extra contribution arises at $r=0$; this is clearly seen when the relation

$$\boldsymbol{\nabla}^2\frac{1}{r}=-4\pi\delta(\boldsymbol{r}) \tag{E.2.30}$$

is used in the second term. However, a similar singular contribution arises from the first term, too, as it was shown in Section 3.3.1. By repeating the derivation presented there, the interaction Hamiltonian can be given as the sum of two terms:

$$\mathcal{H}_{\rm int}=\mathcal{H}_{\rm dipole}+\mathcal{H}_{\rm contact}\,, \tag{E.2.31}$$

where, in addition to the usual dipole–dipole interaction term, the Fermi contact interaction

$$\mathcal{H}_{\rm contact}=-\frac{2\mu_0}{3}\boldsymbol{\mu}_j\cdot\boldsymbol{\mu}_{\rm n}\,\delta(\boldsymbol{r}) \tag{E.2.32}$$

appears. This term is related to the overlap of the two moments, and plays a fundamental role in hyperfine interactions.

In magnetic scattering the spin direction can change for neutrons as well as for electrons bound to atoms. Therefore, besides the wave vector of the neutron and the spatial part of the electron wavefunction, the spin states of the neutron and electron must also be explicitly included in the wavefunction of the initial and final states. Denoting the spatial part of the electron wavefunction before and after scattering by $\phi_{\rm i}$ and $\phi_{\rm f}$, the initial and final states are

$$\begin{aligned}|i,\boldsymbol{k}\rangle &= \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}_{\rm n}}|s_{\rm n}\rangle\phi_{\rm i}(\boldsymbol{r}_j)|\sigma_j\rangle\,,\\ |f,\boldsymbol{k}'\rangle &= \mathrm{e}^{\mathrm{i}\boldsymbol{k}'\cdot\boldsymbol{r}_{\rm n}}|s'_{\rm n}\rangle\phi_{\rm f}(\boldsymbol{r}_j)|\sigma'_j\rangle\,.\end{aligned} \tag{E.2.33}$$

The matrix element between these initial and final states is

$$\left\langle\boldsymbol{k},i\left|\mathcal{H}_{\rm int}\right|f,\boldsymbol{k}'\right\rangle=\left\langle\sigma_j,s_{\rm n}\right|\int\mathrm{d}\boldsymbol{r}_{\rm n}\,\mathrm{d}\boldsymbol{r}_j\phi_{\rm i}^*(\boldsymbol{r}_j)\phi_{\rm f}\left(\boldsymbol{r}_j\right)\mathrm{e}^{-\mathrm{i}(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{r}_{\rm n}}\mathcal{H}_{\rm int}\left|s_{\rm n},\sigma_j\right\rangle. \tag{E.2.34}$$

To evaluate this expression, the form (E.2.28) of the interaction is best used. Integration by parts of the first term gives

$$\begin{aligned}\int\mathrm{d}\boldsymbol{r}\mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}}\left(\boldsymbol{\mu}_j\cdot\boldsymbol{\nabla}\right)\left(\boldsymbol{\mu}_{\rm n}\cdot\boldsymbol{\nabla}\frac{1}{r}\right) &= \mathrm{i}\boldsymbol{K}\cdot\boldsymbol{\mu}_j\int\mathrm{d}\boldsymbol{r}\mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}}\boldsymbol{\mu}_{\rm n}\cdot\boldsymbol{\nabla}\frac{1}{r}\\ &= -\left(\boldsymbol{K}\cdot\boldsymbol{\mu}_j\right)\left(\boldsymbol{K}\cdot\boldsymbol{\mu}_{\rm n}\right)\int\mathrm{d}\boldsymbol{r}\frac{1}{r}\mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}}\\ &= -\left(\boldsymbol{K}\cdot\boldsymbol{\mu}_j\right)\left(\boldsymbol{K}\cdot\boldsymbol{\mu}_{\rm n}\right)\frac{4\pi}{K^2}\,.\end{aligned} \tag{E.2.35}$$

Making use of (E.2.30) in the second term, the matrix element becomes

$$\langle \boldsymbol{k}, i\,|\mathcal{H}_{\rm int}|\,f, \boldsymbol{k}'\rangle = \mu_0 \langle \sigma_j, s_{\rm n}\big|(\boldsymbol{K}\cdot\boldsymbol{\mu}_j)(\boldsymbol{K}\cdot\boldsymbol{\mu}_{\rm n})K^{-2} - \boldsymbol{\mu}_j\cdot\boldsymbol{\mu}_{\rm n}\big|s'_{\rm n}, \sigma'_j\rangle F(\boldsymbol{K})\,, \tag{E.2.36}$$

where $F(\boldsymbol{K})$ is the atomic form factor introduced in (8.1.27),

$$F(\boldsymbol{K}) = \int \mathrm{d}\boldsymbol{r}_j \phi_{\rm i}^*(\boldsymbol{r}_j)\phi_{\rm f}(\boldsymbol{r}_j)\mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}_j}\,. \tag{E.2.37}$$

Expressed in terms of the unit vector $\boldsymbol{e}$ along the direction of $\boldsymbol{K}$, the matrix element can be rewritten as

$$\begin{aligned}\langle \boldsymbol{k}, i\,|\mathcal{H}_{\rm int}|\,f, \boldsymbol{k}'\rangle &= -\mu_0 \langle \sigma_j, s_{\rm n}\,\big|\boldsymbol{\mu}_{\rm n}\cdot\big(\boldsymbol{e}\times\big[\boldsymbol{\mu}_j\times\boldsymbol{e}\big]\big)\big|\,s'_{\rm n}, \sigma'_j\rangle F(\boldsymbol{K}) \\ &= -\mu_0 \langle s_{\rm n}|\boldsymbol{\mu}_{\rm n}|s'_{\rm n}\rangle\cdot\langle\sigma_j|\left(\boldsymbol{e}\times\big[\boldsymbol{\mu}_j\times\boldsymbol{e}\big]\right)|\sigma'_j\rangle F(\boldsymbol{K})\,.\end{aligned} \tag{E.2.38}$$

If z_m electrons of the mth atom participate in the magnetic scattering process, and their common wavefunction is $\Psi_m(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots \boldsymbol{r}_{z_m})$, then the cross section contains the *magnetic form factor*

$$\begin{aligned}F_m(\boldsymbol{K}) = &\int \prod \mathrm{d}\boldsymbol{r}_\nu \Psi_m^*(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots \boldsymbol{r}_{z_m}) \\ &\times \sum_{\nu=1}^{z_m} \frac{\mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}_\nu}(\boldsymbol{s}_\nu\cdot\boldsymbol{S}_m)}{S(S+1)}\Psi_m(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots \boldsymbol{r}_{z_m})\,,\end{aligned} \tag{E.2.39}$$

which characterizes the spatial distribution of "magnetic" electrons. Here $\boldsymbol{S}_m$ is the total spin of the mth atom; the associated magnetic moment $\boldsymbol{\mu}_m$ is given by

$$\boldsymbol{\mu}_m = g\mu_{\rm B}\boldsymbol{S}_m\,. \tag{E.2.40}$$

This total atomic moment appears in the matrix element, too,

$$\langle \boldsymbol{k}, i\,|\mathcal{H}_{\rm int}|\,f, \boldsymbol{k}'\rangle = -\mu_0 \langle s_{\rm n}|\boldsymbol{\mu}_{\rm n}|s'_{\rm n}\rangle\cdot\langle\sigma_j|\left(\boldsymbol{e}\times[\boldsymbol{\mu}_m\times\boldsymbol{e}]\right)|\sigma'_j\rangle F_m(\boldsymbol{K})\,. \tag{E.2.41}$$

In expressions (E.2.37) and (E.2.39) for the atomic form factor $\boldsymbol{r}_j$ and $\boldsymbol{r}_\nu$ denote the absolute position of the electron. Changing to the relative coordinate with respect to the position $\boldsymbol{r}_m$ of the atom, the substitution

$$F_m(\boldsymbol{K}) \to F_m(\boldsymbol{K})\mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}_m} \tag{E.2.42}$$

allows for the separation of a phase factor due to the actual atomic position, and the obtained $F_m(\boldsymbol{K})$ is the true atomic magnetic form factor.

The cross section is obtained by summing over all atoms. The calculations below are for the simple case when the incident neutron beam is unpolarized and the polarization of the scattered beam is not measured, either. By characterizing the probabilities of the possible spin states of the neutron by the density matrix $\rho_{\rm n}$, summing over final-state spins, and using the density matrix ρ for specifying the initial spin state of the atoms, the cross section is given by

$$\begin{aligned}\frac{\mathrm{d}^2\sigma}{\mathrm{d}\Omega\,\mathrm{d}\varepsilon} = \frac{k'}{k}\left(\frac{m_\mathrm{n}\mu_0}{2\pi\hbar^2}\right)^2 \sum_{mn}\sum_{s_\mathrm{n},s'_\mathrm{n}}\sum_{\{\sigma_l\},\{\sigma'_l\}} &\langle s_\mathrm{n}|\rho_\mathrm{n}|s_\mathrm{n}\rangle\langle\{\sigma_l\}|\rho|\{\sigma_l\}\rangle \\ &\times F_m(\boldsymbol{K})\mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}_m}F_n^*(\boldsymbol{K})\mathrm{e}^{\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}_n} \\ &\times \langle s_\mathrm{n}|\boldsymbol{\mu}_\mathrm{n}|s'_\mathrm{n}\rangle\cdot\langle\{\sigma_l\}|\left(\boldsymbol{e}\times[\boldsymbol{S}_m\times\boldsymbol{e}]\right)|\{\sigma'_l\}\rangle \\ &\times \langle s'_\mathrm{n}|\boldsymbol{\mu}_\mathrm{n}|s_\mathrm{n}\rangle\cdot\langle\{\sigma'_l\}|\left(\boldsymbol{e}\times[\boldsymbol{S}_n\times\boldsymbol{e}]\right)|\{\sigma_l\}\rangle\delta(\varepsilon-E_\mathrm{f}+E_\mathrm{i})\,.\end{aligned} \tag{E.2.43}$$

Separating the factors that contain the neutron magnetic moment, and expressing the moment with the neutron g-factor, the nuclear magneton and the spin as

$$\boldsymbol{\mu}_\mathrm{n} = g_\mathrm{n}\mu_\mathrm{N}\boldsymbol{s}_\mathrm{n}\,, \tag{E.2.44}$$

we have

$$\begin{aligned}\sum_{s_\mathrm{n},s'_\mathrm{n}}\langle s_\mathrm{n}|\rho_\mathrm{n}|s_\mathrm{n}\rangle\langle s_\mathrm{n}|\mu_\mathrm{n}^\alpha|s'_\mathrm{n}\rangle\langle s'_\mathrm{n}|\mu_\mathrm{n}^\beta|s_\mathrm{n}\rangle &= \sum_{s_\mathrm{n}}\langle s_\mathrm{n}|\rho_\mathrm{n}|s_\mathrm{n}\rangle\langle s_\mathrm{n}|\mu_\mathrm{n}^\alpha\mu_\mathrm{n}^\beta|s_\mathrm{n}\rangle \\ &= \tfrac{1}{4}\left(g_\mathrm{n}\mu_\mathrm{N}\right)^2\delta_{\alpha\beta}\sum_{s_\mathrm{n}}\langle s_\mathrm{n}|\rho_\mathrm{n}|s_\mathrm{n}\rangle \\ &= \tfrac{1}{4}\left(g_\mathrm{n}\mu_\mathrm{N}\right)^2\delta_{\alpha\beta}\,.\end{aligned} \tag{E.2.45}$$

Now the integral representation can be used for the delta function of energy conservation, and the time dependence can be absorbed in one of the operators, as was done for potential scattering. Then, by making use of the completeness relation to sum over the intermediate spin states of the electron system, by exploiting the identity

$$(\boldsymbol{e}\times[\boldsymbol{a}\times\boldsymbol{e}])\cdot(\boldsymbol{e}\times[\boldsymbol{b}\times\boldsymbol{e}]) = \boldsymbol{a}\cdot\boldsymbol{b} - (\boldsymbol{a}\cdot\boldsymbol{e})\,(\boldsymbol{b}\cdot\boldsymbol{e})\,, \tag{E.2.46}$$

and finally by eliminating the atomic moments in favor of the spin variables, one obtains

$$\begin{aligned}\frac{\mathrm{d}^2\sigma}{\mathrm{d}\Omega\,\mathrm{d}\varepsilon} = \tfrac{1}{4}\left(g_\mathrm{n}\mu_\mathrm{N}\right)^2\left(g\mu_\mathrm{B}\right)^2\frac{k'}{k}\left(\frac{m_\mathrm{n}\mu_0}{2\pi\hbar^2}\right)^2 &\sum_{mn,\alpha\beta}\left(\delta_{\alpha\beta}-e_\alpha e_\beta\right) \\ &\times F_m(\boldsymbol{K})F_n^*(\boldsymbol{K})\mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}_m}\mathrm{e}^{\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{r}_n} \\ &\times\frac{1}{2\pi\hbar}\int_{-\infty}^{\infty}\mathrm{d}t\,\mathrm{e}^{\mathrm{i}\varepsilon t/\hbar}\left\langle S_m^\alpha(t)S_n^\beta(0)\right\rangle.\end{aligned} \tag{E.2.47}$$

Atomic motions were ignored in the previous derivation. When due care is taken of them, atomic positions also appear in the thermal averages, leading to

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}\Omega\,\mathrm{d}\varepsilon} = \tfrac{1}{4}\left(g_\mathrm{n}\mu_\mathrm{N}\right)^2\left(g\mu_\mathrm{B}\right)^2\frac{k'}{k}\left(\frac{m_\mathrm{n}\mu_0}{2\pi\hbar^2}\right)^2\sum_{mn,\alpha\beta}(\delta_{\alpha\beta}-e_\alpha e_\beta)$$
$$\times F_m(\boldsymbol{K})F_n^*(\boldsymbol{K})\mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{R}_m}\mathrm{e}^{\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{R}_n} \qquad \text{(E.2.48)}$$
$$\times\frac{1}{2\pi\hbar}\int\limits_{-\infty}^{\infty}\mathrm{d}t\,\mathrm{e}^{\mathrm{i}\varepsilon t/\hbar}\left\langle\mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{u}_m(t)}S_m^\alpha(t)\mathrm{e}^{\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{u}_n}S_n^\beta(0)\right\rangle.$$

If spin–phonon interactions are neglected, the averaging procedures can be performed separately for spins and atomic positions. When calculating the contribution of magnetic scattering, atoms are assumed to occupy their equilibrium positions, and so

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}\Omega\,\mathrm{d}\varepsilon} = \tfrac{1}{4}\left(g_\mathrm{n}\mu_\mathrm{N}\right)^2\left(g\mu_\mathrm{B}\right)^2\frac{k'}{k}\left(\frac{m_\mathrm{n}\mu_0}{2\pi\hbar^2}\right)^2\sum_{mn}F_m(\boldsymbol{K})F_n^*(\boldsymbol{K})\mathrm{e}^{-\mathrm{i}\boldsymbol{K}\cdot(\boldsymbol{R}_m-\boldsymbol{R}_n)}$$
$$\times\sum_{\alpha\beta}(\delta_{\alpha\beta}-e_\alpha e_\beta)\frac{1}{2\pi\hbar}\int\limits_{-\infty}^{\infty}\mathrm{d}t\,\mathrm{e}^{\mathrm{i}\varepsilon t/\hbar}\left\langle S_m^\alpha(t)S_n^\beta(0)\right\rangle. \qquad \text{(E.2.49)}$$

The spatial distribution of the moment is characterized by the atomic magnetic form factor, while the spin correlation function describes spin dynamics.

References

1. E. Balcar and S. W. Lovesey, *Theory of Magnetic Neutron and Photon Scattering*, Oxford University Press, Oxford (1989).
2. S. W. Lovesey, *Theory of Neutron Scattering from Condensed Matter*, Vol. I: *Nuclear Scattering*, Clarendon Press, Oxford (1986).
3. S. W. Lovesey, *Theory of Neutron Scattering from Condensed Matter*, Vol. II: *Polarization Effects and Magnetic Scattering*, Clarendon Press, Oxford (1986).
4. W. Marshall and S. W. Lovesey, *Theory of Thermal Neutron Scattering*, Clarendon Press, Oxford (1971).

F

The Algebra of Angular-Momentum and Spin Operators

In this Appendix we shall first list some basic relations for quantum mechanical angular-momentum operators, and then turn to the special properties of spin operators.

F.1 Angular Momentum

It was mentioned in Appendix D that the dimensionless orbital angular momentum operator defined via $\hbar \boldsymbol{L} = \boldsymbol{r} \times \boldsymbol{p}$ is the generator of the transformations of the continuous rotation group. We shall discuss this relation in more detail below, and then present the irreducible representations of the rotation group as well as the addition theorem for angular momenta.

F.1.1 Angular Momentum and the Rotation Group

The set of all possible rotations around any axis through the origin form a continuous group, the full rotation group SO(3). The infinitesimal rotation through angle $\delta\alpha$ around the axis characterized by the unit vector $\boldsymbol{n}$ is denoted by $C_{\boldsymbol{n}}(\delta\alpha)$; in this operation the end point of the vector $\boldsymbol{r}$ is displaced by

$$\delta \boldsymbol{r} = \delta\alpha\, \boldsymbol{n} \times \boldsymbol{r}\,. \tag{F.1.1}$$

In the space of the functions $\psi(\boldsymbol{r})$ an operator $O\big(C_{\boldsymbol{n}}(\delta\alpha)\big)$ is associated with this transformation; it takes function $\psi(\boldsymbol{r})$ into

$$\begin{aligned} O\big(C_{\boldsymbol{n}}(\delta\alpha)\big)\psi(\boldsymbol{r}) &= \psi\big(C_{\boldsymbol{n}}^{-1}(\delta\alpha)\boldsymbol{r}\big) = \psi(\boldsymbol{r} - \delta\boldsymbol{r}) \\ &= [1 - \delta\alpha\, \boldsymbol{n} \cdot (\boldsymbol{r} \times \boldsymbol{\nabla})]\psi(\boldsymbol{r})\,, \end{aligned} \tag{F.1.2}$$

that is, the rotation operator can be written as

$$O\big(C_{\boldsymbol{n}}(\delta\alpha)\big) = 1 - \delta\alpha\, \boldsymbol{n} \cdot (\boldsymbol{r} \times \boldsymbol{\nabla}) = 1 - \mathrm{i}\delta\alpha(\boldsymbol{n} \cdot \boldsymbol{L})\,. \tag{F.1.3}$$

It then follows that for a rotation through an arbitrary angle α

$$O\big(C_{\boldsymbol{n}}(\alpha)\big) = \mathrm{e}^{-\mathrm{i}\alpha(\boldsymbol{n}\cdot\boldsymbol{L})}\,. \tag{F.1.4}$$

For conciseness we shall use the commoner notation

$$R_{\boldsymbol{n}}(\alpha) \equiv O\big(C_{\boldsymbol{n}}(\alpha)\big)\,. \tag{F.1.5}$$

The components of the operator $\boldsymbol{L}$ satisfy the commutation relations

$$[L_x, L_y] = \mathrm{i}L_z\,, \quad [L_y, L_z] = \mathrm{i}L_x\,, \quad [L_z, L_x] = \mathrm{i}L_y\,, \tag{F.1.6}$$

which can be concisely written as

$$[L_\alpha, L_\beta] = \mathrm{i}\sum_\gamma \varepsilon_{\alpha\beta\gamma} L_\gamma\,, \tag{F.1.7}$$

or as

$$\boldsymbol{L}\times\boldsymbol{L} = \mathrm{i}\boldsymbol{L}\,, \tag{F.1.8}$$

where $\varepsilon_{\alpha\beta\gamma}$ is the Levi-Civita tensor. We shall call any operator $\boldsymbol{J}$ an angular-momentum operator if it appears as a generator for rotations and its components satisfy the commutation relation

$$[J_\alpha, J_\beta] = \mathrm{i}\sum_\gamma \varepsilon_{\alpha\beta\gamma} J_\gamma\,, \tag{F.1.9}$$

even if the transformation does not concern a complex function $\psi(\boldsymbol{r})$ but a spinor or a vector function.

It is not always convenient to give the rotation operator in this form. Namely, if the rotation axis $\boldsymbol{n}$ is specified by its polar angles θ and φ, the term

$$-\mathrm{i}\alpha(\sin\theta\cos\varphi J_x + \sin\theta\sin\varphi J_y + \cos\theta J_z) \tag{F.1.10}$$

appears in the exponent – and since the components of the angular-momentum operator do not commute, this expression is more difficult to evaluate. Therefore rotations are customarily characterized using the Euler angles. Several widely used conventions exist; we shall adopt the one in which a rotation through ϕ around the z-axis is followed by a rotation through ϑ around the new y'-axis, and finally by a rotation through ψ around the even newer z''-axis:

$$R_{\boldsymbol{n}}(\alpha) = R(\phi,\vartheta,\psi) = R_{z''}(\psi)R_{y'}(\vartheta)R_z(\phi)\,. \tag{F.1.11}$$

In terms of the generators of rotations,

$$R(\phi,\vartheta,\psi) = \mathrm{e}^{-\mathrm{i}\psi J_{z''}}\mathrm{e}^{-\mathrm{i}\vartheta J_{y'}}\mathrm{e}^{-\mathrm{i}\phi J_z}. \tag{F.1.12}$$

An even simpler form is obtained when it is recognized that the y'-axis is obtained from the y-axis via a rotation around the z-axis through angle ϕ, that is

$$J_{y'} = R_z(\phi) J_y R_z^{-1}(\phi)\,, \tag{F.1.13}$$

and therefore

$$\mathrm{e}^{-\mathrm{i}\vartheta J_{y'}} = \mathrm{e}^{-\mathrm{i}\phi J_z} \mathrm{e}^{-\mathrm{i}\vartheta J_y} \mathrm{e}^{\mathrm{i}\phi J_z}. \tag{F.1.14}$$

Using the corresponding expression for the operator of the rotation around the z''-axis,

$$R(\phi, \vartheta, \psi) = \mathrm{e}^{-\mathrm{i}\phi J_z} \mathrm{e}^{-\mathrm{i}\vartheta J_y} \mathrm{e}^{-\mathrm{i}\psi J_z} \tag{F.1.15}$$

is finally obtained.

A possible representation of the rotation matrices can be easily obtained from determining their action on the function $\psi(\boldsymbol{r}) = \boldsymbol{r}$. For a rotation around the z-axis through angle ϕ,

$$\mathsf{R}_z(\phi) = \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{F.1.16}$$

Likewise, for a rotation around the y-axis through angle ϑ,

$$\mathsf{R}_y(\vartheta) = \begin{pmatrix} \cos\vartheta & 0 & -\sin\vartheta \\ 0 & 1 & 0 \\ \sin\vartheta & 0 & \cos\vartheta \end{pmatrix}, \tag{F.1.17}$$

while for a rotation around the z-axis through angle ψ,

$$\mathsf{R}_z(\psi) = \begin{pmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{F.1.18}$$

Owing to the transformations of the axes, rotations appear in the reverse order in (F.1.15):

$$\mathsf{R}(\phi, \vartheta, \psi) = \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\vartheta & 0 & -\sin\vartheta \\ 0 & 1 & 0 \\ \sin\vartheta & 0 & \cos\vartheta \end{pmatrix} \begin{pmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{F.1.19}$$

Thus we have obtained the representation of the rotation group by 3×3 orthogonal matrices of unit determinant, as $\mathsf{R}^{\mathrm{T}} = \mathsf{R}^{-1}$. Spatial rotations in three dimensions are indeed isomorphic to the group SO(3).

F.1.2 The Irreducible Representations of the Rotation Group

To determine further representations of the rotation group, we shall exploit the fact that $\boldsymbol{J}^2$ and J_z commute, and therefore have a common set of eigenfunctions. Denoting the eigenvalues of $\boldsymbol{J}^2$ by $j(j+1)$ and those of J_z by m,

it follows from the properties of the operators that j can take only positive, while m any integer or half-integer[1] value,

$$\begin{aligned} j &= 0, 1/2, 1, 3/2, 2, \ldots, \\ m &= 0, \pm 1/2, \pm 1, \pm 3/2, \pm 2, \ldots, \end{aligned} \tag{F.1.20}$$

but for a given j the eigenvalues of m can only be

$$m = -j, -j+1, \ldots, j\,. \tag{F.1.21}$$

The angular-momentum operator can be given in the space of states with quantum numbers j, m – denoted as $|j, m\rangle$ – by $(2j+1)\times(2j+1)$ matrices. To determine the matrix element, the raising and lowering operators $J_\pm = J_x \pm \mathrm{i}J_y$ are introduced; they increase and decrease the component J_z by unity:

$$\begin{aligned} J_+|j,m\rangle &= [j(j+1) - m(m+1)]^{1/2}|j, m+1\rangle \\ &= (j-m)^{1/2}(j+m+1)^{1/2}|j, m+1\rangle\,, \\ J_-|j,m\rangle &= [j(j+1) - m(m-1)]^{1/2}|j, m-1\rangle \\ &= (j+m)^{1/2}(j-m+1)^{1/2}|j, m+1\rangle\,, \end{aligned} \tag{F.1.22}$$

in line with the relations

$$\begin{aligned} J_+J_- &= \boldsymbol{J}^2 - J_z(J_z-1)\,, \\ J_-J_+ &= \boldsymbol{J}^2 - J_z(J_z+1)\,, \\ \boldsymbol{J}^2 &= J_z^2 + \tfrac{1}{2}(J_+J_- + J_-J_+)\,. \end{aligned} \tag{F.1.23}$$

For $j = 1/2$ the generators are 2×2 matrices of the form

$$J_x = \tfrac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad J_y = \tfrac{1}{2}\begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \qquad J_z = \tfrac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{F.1.24}$$

while for $j = 1$

$$J_x = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad J_y = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & -\mathrm{i} & 0 \\ \mathrm{i} & 0 & -\mathrm{i} \\ 0 & \mathrm{i} & 0 \end{pmatrix}, \quad J_z = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \tag{F.1.25}$$

It is straightforward to show that a suitable unitary transformation can take these matrices into (D.1.72), that is, the two forms are equivalent.

Expressing the rotation matrices in terms of these generators, as in (F.1.15), the matrix elements in the space of states $|j, m\rangle$ are

$$R^j_{mm'}(\phi, \vartheta, \psi) = \left\langle j, m \middle| \mathrm{e}^{-\mathrm{i}\phi J_z}\mathrm{e}^{-\mathrm{i}\vartheta J_y}\mathrm{e}^{-\mathrm{i}\psi J_z} \middle| j, m' \right\rangle. \tag{F.1.26}$$

[1] By half-integer or half-odd-integer numbers we mean numbers of the form $(2n+1)/2$, where n is integer.

It can be shown that these representations are irreducible, and a single irreducible representation is associated with each j. The irreducible representations can thus be characterized by the quantum number j, so in what follows we shall use the customary notation $\mathsf{D}_j(\phi, \vartheta, \psi)$.

To determine the character of the irreducible representation of quantum number j associated with a rotation through angle φ, we may choose the rotation axis along the direction of the z-axis of quantization, as rotations through identical angles around different axes are conjugate to each other. On the space of the states $|j, m\rangle$ the matrix of the generator of infinitesimal rotations is diagonal:

$$J_z = \begin{pmatrix} j & 0 & \dots & 0 \\ 0 & j-1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & -j \end{pmatrix} . \tag{F.1.27}$$

This implies directly that the matrix for the rotation through angle ϕ around the quantization axis is

$$\mathsf{D}_j(\phi) = \mathrm{e}^{-\mathrm{i}\phi J_z} = \begin{pmatrix} \mathrm{e}^{-\mathrm{i}j\phi} & 0 & \dots & 0 \\ 0 & \mathrm{e}^{-\mathrm{i}(j-1)\phi} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mathrm{e}^{\mathrm{i}j\phi} \end{pmatrix} , \tag{F.1.28}$$

consequently the character of this irreducible representation is

$$\chi_j(\phi) = \sum_{m=-j}^{j} \mathrm{e}^{\mathrm{i}m\phi} = \frac{\sin(j+\frac{1}{2})\phi}{\sin\frac{1}{2}\phi} . \tag{F.1.29}$$

The characters satisfy the orthogonality relation

$$\frac{1}{\pi} \int_0^\pi \mathrm{d}\phi (1-\cos\phi) \chi_j(\phi) \chi_{j'}(\phi) = \delta_{jj'} . \tag{F.1.30}$$

F.1.3 Orbital Angular Momentum and Spin

A rotations through 2π around the z-axis takes the state $|j, m\rangle$ into

$$R_z(2\pi)|j, m\rangle = \mathrm{e}^{-\mathrm{i}2\pi m}|j, m\rangle . \tag{F.1.31}$$

As m is an integer, the phase factor is unity for integer values of j and -1 for half-integer j. This means that integer js are associated with single-valued representations, and half-integer js with double-valued ones. Since a wavefunction that depends only on the position variable $\boldsymbol{r}$ must be single-valued,

the orbital angular momentum quantum number (henceforth denoted by l) can take only integer values in such states.

On the other hand, the wavefunction is a two-component spinor for electrons:

$$\psi(\boldsymbol{r},\sigma)=\begin{pmatrix}\psi_+(\boldsymbol{r})\\ \psi_-(\boldsymbol{r})\end{pmatrix}. \tag{F.1.32}$$

Now, in addition to transforming the position coordinate $\boldsymbol{r}$, the operator $R(\phi,\vartheta,\psi)$ also mixes the spinor components; as the spinor has two components ($j=1/2$), the transformation is specified by the 2×2 matrix $\mathsf{D}_{1/2}(\phi,\vartheta,\psi)$:

$$R(\phi,\vartheta,\psi)\psi_\sigma(\boldsymbol{r})=\sum_{\sigma'}D^{1/2}_{\sigma'\sigma}(\phi,\vartheta,\psi)\psi_{\sigma'}(R^{-1}\boldsymbol{r})\,, \tag{F.1.33}$$

as given in (D.1.45). This wavefunction is taken into itself only by rotations through 4π.

An infinitesimal rotation around the z-axis gives

$$R_z(\delta\phi)\begin{pmatrix}\psi_+(\boldsymbol{r})\\ \psi_-(\boldsymbol{r})\end{pmatrix}=\big[1-\mathrm{i}\delta\phi\big(s_z+l_z\big)\big]\begin{pmatrix}\psi_+(\boldsymbol{r})\\ \psi_-(\boldsymbol{r})\end{pmatrix}, \tag{F.1.34}$$

where s_z and l_z are the generators for the rotation of the spinor components and the spatial part of the wavefunction. Similar formulas apply to rotations around the other axes. Denoting the dimensionless angular momentum associated with spin and with orbital angular momentum by $\boldsymbol{s}$ and $\boldsymbol{l}$, the total dimensionless angular momentum of the electron is

$$\boldsymbol{j}=\boldsymbol{s}+\boldsymbol{l}\,. \tag{F.1.35}$$

For particles described by vector fields the vector components are transformed according to the representation D_1 of the rotation group ($j=1$), thus they can be considered as spin-1 particles ($s=1$). Spin can therefore take integer or half-integer values.

F.1.4 Addition Theorem for Angular Momenta

The combination of two systems, of angular momenta $\boldsymbol{j}_1$ and $\boldsymbol{j}_2$ leads to a net angular momentum of $\boldsymbol{j}=\boldsymbol{j}_1+\boldsymbol{j}_2$. The eigenvalue of its square is $j(j+1)$, while that of its z component is m, where

$$j=|j_1-j_2|,\ldots,j_1+j_2;\qquad m=-j,\ldots,j. \tag{F.1.36}$$

Now consider the eigenstates $|j_1m_1\rangle$ and $|j_2m_2\rangle$ of the pairs of operators ($\boldsymbol{j}_1^2$, j_{1z}) and ($\boldsymbol{j}_2^2$, j_{2z}) with quantum numbers j_1,m_1 and j_2,m_2; their products span a $(2j_1+1)(2j_2+1)$-dimensional space. Counting the eigenstates $|jm\rangle$ using the allowed quantum numbers j and m of the operators $\boldsymbol{j}^2$ and j_z leads to the same result, as

$$\sum_{j=|j_1-j_2|}^{j_1+j_2}(2j+1)=(2j_1+1)(2j_2+1)\,. \tag{F.1.37}$$

This means that the direct product of two representations (of indices j_1 and j_2) can be reduced, and this reduction leads to the direct sum of the representations of indices ranging from $|j_1-j_2|$ to j_1+j_2, each of them occurring once:

$$D_{j_1}\otimes D_{j_2}=\sum_{\mu=|j_1-j_2|}^{j_1+j_2}D_\mu\,. \tag{F.1.38}$$

Writing these states as the linear combination of the products of states $|j_1m_1\rangle$ and $|j_2m_2\rangle$, the arising coefficients are the Clebsch–Gordan coefficients $(j_1j_2m_1m_2|jm)$ or the Wigner $3j$ symbols defined as

$$\begin{pmatrix} j_1 & j_2 & j\\ m_1 & m_2 & -m\end{pmatrix}=\frac{(-1)^{j_1-j_2+m}}{(2j+1)^{1/2}}(j_1j_2m_1m_2|jm)\,, \tag{F.1.39}$$

in terms of which

$$\begin{aligned}|jm\rangle&=\sum_{m_1m_2}(j_1j_2m_1m_2|jm)|j_1m_1\rangle|j_2m_2\rangle \qquad\text{(F.1.40)}\\ &=\sum_{m_1m_2}(-1)^{j_1-j_2+m}(2j+1)^{1/2}\begin{pmatrix} j_1 & j_2 & j\\ m_1 & m_2 & -m\end{pmatrix}|j_1m_1\rangle|j_2m_2\rangle\,.\end{aligned}$$

This expression vanishes unless $m=m_1+m_2$. Various notations are used in the literature for the Clebsch–Gordan coefficients, for example,

$$C^{jm}_{j_1m_1j_2m_2}=(j_1j_2m_1m_2|jm)\,. \tag{F.1.41}$$

They can be chosen as real, and can be given explicitly as

$$\begin{aligned}(j_1j_2m_1m_2|jm)&=\sqrt{\frac{(2j+1)(j_1+j_2-j)!(j+j_1-j_2)!(j+j_2-j_1)!}{(j_1+j_2+j+1)!}}\\ &\times\delta_{m,m_1+m_2}\sqrt{(j_1+m_1)!(j_1-m_1)!(j_2+m_2)!(j_2-m_2)!(j+m)!(j-m)!}\\ &\times\sum_k(-1)^k\big[k!(j_1+j_2-j-k)!(j_1-m_1-k)!(j_2+m_2-k)!\\ &\qquad\times(j-j_2+m_1+k)!(j-j_1-m_2+k)!\big]^{-1}\,,\end{aligned} \tag{F.1.42}$$

where the restriction on the k-sum is imposed by the requirement that the argument of the factorial function should be nonnegative.

These coefficients satisfy the orthogonality relations

$$\begin{aligned}\sum_{m_1=-j_1}^{j_1}\sum_{m_2=-j_2}^{j_2}(j_1j_2m_1m_2|jm)(j_1j_2m_1m_2|j'm')&=\delta_{jj'}\delta_{mm'}\,,\\ \sum_{j=|j_1-j_2|}^{j_1+j_2}\sum_{m=-j}^{j}(j_1j_2m_1m_2|jm)(j_1j_2m_1'm_2'|jm)&=\delta_{m_1m_1'}\delta_{m_2m_2'}\,.\end{aligned} \tag{F.1.43}$$

For the $3j$ symbols these relations are formulated as

$$\begin{aligned}\sum_{m_1 m_2}(2j+1)\begin{pmatrix} j_1 & j_2 & j \\ m_1 & m_2 & m\end{pmatrix}\begin{pmatrix} j_1 & j_2 & j' \\ m_1 & m_2 & m'\end{pmatrix} &= \delta_{jj'}\delta_{mm'}\,,\\ \sum_{jm}(2j+1)\begin{pmatrix} j_1 & j_2 & j \\ m_1 & m_2 & m\end{pmatrix}\begin{pmatrix} j_1 & j_2 & j \\ m_1' & m_2' & m\end{pmatrix} &= \delta_{m_1 m_1'}\delta_{m_2 m_2'}\,.\end{aligned} \tag{F.1.44}$$

It is then directly established that

$$|j_1 m_1\rangle|j_2 m_2\rangle = \sum_{j=|j_1-j_2|}^{j_1+j_2}\sum_{m=-j}^{j}(j_1 j_2 m_1 m_2|jm)|jm\rangle\,. \tag{F.1.45}$$

Using the spherical harmonics as eigenfunctions for integer values of j_1 and j_2, and noting that they are normalized over the surface of the unit sphere, the previous expression implies the addition theorem for spherical harmonics, (C.4.30).

The sum of more than two angular momenta involves the Wigner $6j$ and $9j$ symbols or the Racah coefficients.[2]

F.2 Orbital Angular Momentum

It follows from the explicit form $\hbar\boldsymbol{l} = \boldsymbol{r}\times\boldsymbol{p}$ of the orbital angular momentum that the corresponding operator satisfies the following commutation relations with the position and momentum operators:

$$\begin{aligned}[l_\alpha, x_\beta] &= \mathrm{i}\sum_\gamma \varepsilon_{\alpha\beta\gamma}x_\gamma\,,\\ [l_\alpha, p_\beta] &= \mathrm{i}\sum_\gamma \varepsilon_{\alpha\beta\gamma}p_\gamma\,.\end{aligned} \tag{F.2.1}$$

For orbital angular momentum the quantum number l can take nonnegative integer values. It can be shown that a single eigenstate belongs to each value of l and m. This is obtained most easily from the spherical coordinate form of the angular momentum operator:

$$L_\pm = L_x \pm \mathrm{i}L_y = \mathrm{e}^{\pm\mathrm{i}\varphi}\left[\pm\frac{\partial}{\partial\theta} + \mathrm{i}\cot\theta\frac{\partial}{\partial\varphi}\right], \qquad L_z = \frac{1}{\mathrm{i}}\frac{\partial}{\partial\varphi}\,. \tag{F.2.2}$$

The square of the angular momentum is then

$$L^2 = -\left[\frac{\partial^2}{\partial\theta^2} + \cot\theta\frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right]. \tag{F.2.3}$$

[2] G. Racah, 1942.

Since spherical harmonics satisfy the equation

$$\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}Y_l^m(\theta,\varphi)\right)+\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}Y_l^m(\theta,\varphi)+l(l+1)Y_l^m(\theta,\varphi)=0\,, \tag{F.2.4}$$

they are eigenfunctions of the squared angular momentum operator with an eigenvalue $l(l+1)$, and of the z component of the angular momentum operator with an eigenvalue m. The normalization of the functions is such that the recursion relation

$$L_\pm Y_l^m(\theta,\varphi)=\left[l(l+1)-m(m\pm1)\right]^{1/2}Y_l^{m\pm1}(\theta,\varphi) \tag{F.2.5}$$

holds.

Being the basis functions of the irreducible representations of the rotation group, spherical harmonics are transformed by rotation R into

$$R(\phi,\vartheta,\psi)Y_l^m(\theta,\varphi)=\sum_{m'}Y_l^{m'}(\theta,\varphi)D^l_{m'm}(\phi,\vartheta,\psi)\,. \tag{F.2.6}$$

On the other hand, any single-valued representation over the space of functions of the variables (x,y,z) can be written as the linear combination of those representations D_l that belong to integer indices l, since the latter constitute a complete system.

F.3 Spin Operators

As spin operators play a special role in the description of magnetic properties, it is useful to recall some important relations among them.

F.3.1 Two-Dimensional Representations of the Rotation Group

Using the representation of the generators of the rotation group by the 2×2 matrices given in (F.1.24),

$$R_z(\phi)=\mathrm{e}^{-\mathrm{i}\phi J_z}=1+\sum_{n=1}^{\infty}\frac{(-\mathrm{i}\phi J_z)^n}{n!}=\begin{pmatrix}\mathrm{e}^{-\mathrm{i}\phi/2}&0\\0&\mathrm{e}^{\mathrm{i}\phi/2}\end{pmatrix}, \tag{F.3.1}$$

and

$$R_y(\vartheta)=\mathrm{e}^{-\mathrm{i}\vartheta J_y}=1+\sum_{n=1}^{\infty}\frac{(-\mathrm{i}\vartheta J_y)^n}{n!}=\begin{pmatrix}\cos\frac{1}{2}\vartheta&-\sin\frac{1}{2}\vartheta\\\sin\frac{1}{2}\vartheta&\cos\frac{1}{2}\vartheta\end{pmatrix}. \tag{F.3.2}$$

Thus for the rotation characterized by the Euler angles ϕ, ϑ, and ψ,

$$D_{1/2}=\begin{pmatrix}\cos\frac{1}{2}\vartheta\exp(-\frac{1}{2}\mathrm{i}(\phi+\psi))&-\sin\frac{1}{2}\vartheta\exp(-\frac{1}{2}\mathrm{i}(\phi-\psi))\\\sin\frac{1}{2}\vartheta\exp(\frac{1}{2}\mathrm{i}(\phi-\psi))&\cos\frac{1}{2}\vartheta\exp(\frac{1}{2}\mathrm{i}(\phi+\psi))\end{pmatrix}. \tag{F.3.3}$$

These matrices are in fact unimodular[3] unitary matrices, $D^\dagger = D^{-1}$, indicating that the rotational group of the three-dimensional space, SO(3), can be mapped onto the group SU(2). Some care must nevertheless be exercised as two different matrices in SU(2) correspond to each element in SO(3). If the rotations are complemented by the element $\overline{E}$ of rotation through 2π, the obtained double group SO′(3) will be isomorphic to SU(2).

If instead of the Euler angles the rotation is characterized by the unit vector $\boldsymbol{n}$ of the rotation axis (or by its polar angle) and by the angle of rotation, the expression

$$R_{\boldsymbol{n}}(\alpha) = \mathrm{e}^{-\mathrm{i}\alpha \boldsymbol{n}\cdot\boldsymbol{J}} \tag{F.3.4}$$

simplifies to

$$\begin{aligned} R_{\boldsymbol{n}}(\alpha) &= \cos\tfrac{1}{2}\alpha - \mathrm{i}\boldsymbol{n}\cdot\boldsymbol{J}\sin\tfrac{1}{2}\alpha \\ &= \begin{pmatrix} \cos\frac{1}{2}\alpha - \mathrm{i}\cos\theta\sin\frac{1}{2}\alpha & -\mathrm{i}\sin\theta\mathrm{e}^{-\mathrm{i}\varphi}\sin\frac{1}{2}\alpha \\ -\mathrm{i}\sin\theta\mathrm{e}^{\mathrm{i}\varphi}\sin\frac{1}{2}\alpha & \cos\frac{1}{2}\alpha + \mathrm{i}\cos\theta\sin\frac{1}{2}\alpha \end{pmatrix} \end{aligned} \tag{F.3.5}$$

on account of the simple multiplication rule of the spin-half operators J_x, J_y, and J_z.

F.3.2 Spin Algebra

Naturally, the dimensionless spin operators satisfy the same commutation relations as the orbital angular momentum components:

$$[S_x, S_y] = \mathrm{i}S_z\,, \quad [S_y, S_z] = \mathrm{i}S_x\,, \quad [S_z, S_x] = \mathrm{i}S_y\,. \tag{F.3.6}$$

The eigenvalues of $\boldsymbol{S}^2$ and S_z are $S(S+1)$ and $M = -S, \ldots, S$. However S is not necessarily an integer now: it may be a half-integer as well.

The raising and lowering operators $S_\pm = S_x \pm \mathrm{i}S_y$ are commonly used for spin, too; their commutation relations are

$$[S_z, S_\pm] = \pm S_\pm\,, \qquad [S_+, S_-] = 2S_z\,. \tag{F.3.7}$$

They transform the state $|S, M\rangle$ according to

$$S_\pm|S, M\rangle = \big[S(S+1) - M(M\pm 1)\big]^{1/2}|S, M\pm 1\rangle\,. \tag{F.3.8}$$

For $S = 1/2$ the spin operator can be expressed in terms of the Pauli matrices:

$$S_\alpha = \tfrac{1}{2}\sigma_\alpha\,, \tag{F.3.9}$$

where

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{F.3.10}$$

[3] A matrix is called unimodular if its determinant is unity.

It is readily seen from these formulas that

$$\begin{aligned} \sigma_x^2 = \sigma_y^2 = \sigma_z^2 = 1\,, \\ \sigma_x\sigma_y = -\sigma_y\sigma_x = \mathrm{i}\sigma_z\,, \\ \sigma_y\sigma_z = -\sigma_z\sigma_y = \mathrm{i}\sigma_x\,, \\ \sigma_z\sigma_x = -\sigma_x\sigma_z = \mathrm{i}\sigma_y\,, \end{aligned} \tag{F.3.11}$$

and that for arbitrary vector operators $\boldsymbol{A}$ and $\boldsymbol{B}$

$$(\boldsymbol{\sigma}\cdot\boldsymbol{A})(\boldsymbol{\sigma}\cdot\boldsymbol{B}) = (\boldsymbol{A}\cdot\boldsymbol{B}) + \mathrm{i}\boldsymbol{\sigma}\cdot[\boldsymbol{A}\times\boldsymbol{B}]\,, \tag{F.3.12}$$

where the order of the operators is not immaterial. If $\boldsymbol{A}$ and $\boldsymbol{B}$ are spin operators of magnitude S, then

$$\sum_{\beta\eta}(\boldsymbol{\sigma}_{\alpha\beta}\cdot\boldsymbol{S}_{\varepsilon\eta})(\boldsymbol{\sigma}_{\beta\gamma}\cdot\boldsymbol{S}_{\eta\zeta}) = S(S+1)\delta_{\alpha\gamma}\delta_{\varepsilon\zeta} - (\boldsymbol{\sigma}_{\alpha\gamma}\cdot\boldsymbol{S}_{\varepsilon\zeta})\,. \tag{F.3.13}$$

Likewise,

$$\sum_{\beta\eta}(\boldsymbol{\sigma}_{\alpha\beta}\cdot\boldsymbol{S}_{\eta\zeta})(\boldsymbol{\sigma}_{\beta\gamma}\cdot\boldsymbol{S}_{\varepsilon\eta}) = S(S+1)\delta_{\alpha\gamma}\delta_{\varepsilon\zeta} + (\boldsymbol{\sigma}_{\alpha\gamma}\cdot\boldsymbol{S}_{\varepsilon\zeta}) \tag{F.3.14}$$

and

$$\sum_{\varepsilon\eta}(\boldsymbol{\sigma}_{\alpha\beta}\cdot\boldsymbol{S}_{\varepsilon\eta})(\boldsymbol{\sigma}_{\gamma\delta}\cdot\boldsymbol{S}_{\eta\varepsilon}) = \tfrac{1}{3}S(S+1)(2S+1)(\boldsymbol{\sigma}_{\alpha\beta}\cdot\boldsymbol{\sigma}_{\gamma\delta})\,. \tag{F.3.15}$$

Taking two spin-half operators, the only SU(2)-invariant combination is their product $\boldsymbol{S}_1\cdot\boldsymbol{S}_2$. Higher powers of this product do not give new terms as

$$(\boldsymbol{S}_1\cdot\boldsymbol{S}_2)^2 = \tfrac{3}{16} - \tfrac{1}{2}\boldsymbol{S}_1\cdot\boldsymbol{S}_2\,. \tag{F.3.16}$$

For $S=1$ the customary matrix form of the spin components is

$$S_x = \frac{1}{\sqrt{2}}\begin{pmatrix} 0&1&0\\ 1&0&1\\ 0&1&0 \end{pmatrix},\quad S_y = \frac{1}{\sqrt{2}}\begin{pmatrix} 0&-\mathrm{i}&0\\ \mathrm{i}&0&-\mathrm{i}\\ 0&\mathrm{i}&0 \end{pmatrix},\quad S_z = \begin{pmatrix} 1&0&0\\ 0&0&0\\ 0&0&-1 \end{pmatrix}. \tag{F.3.17}$$

In addition to the product $\boldsymbol{S}_1\cdot\boldsymbol{S}_2$, which is invariant under SU(2) transformations for any S, $(\boldsymbol{S}_1\cdot\boldsymbol{S}_2)^2$ also appears as an independent term in the $S=1$ case – however, higher powers can be expressed in terms of this:

$$(\boldsymbol{S}_1\cdot\boldsymbol{S}_2)^3 = 2 + \boldsymbol{S}_1\cdot\boldsymbol{S}_2 - 2(\boldsymbol{S}_1\cdot\boldsymbol{S}_2)^2\,. \tag{F.3.18}$$

It is fairly straightforward to show that the combination

$$\boldsymbol{S}_1\cdot\boldsymbol{S}_2 + (\boldsymbol{S}_1\cdot\boldsymbol{S}_2)^2 \tag{F.3.19}$$

of the independent quantities $\boldsymbol{S}_1\cdot\boldsymbol{S}_2$ and $(\boldsymbol{S}_1\cdot\boldsymbol{S}_2)^2$ is invariant even under SU(3) transformations for $S=1$. To this end, the spin operators have to be represented by their 3×3 matrices, and the latter have to be expressed with the Gell-Mann matrices.

F.3.3 Projection Operators

We have seen for angular momentum in general that the magnitude of the resultant of two angular momenta, $\boldsymbol{j}_1$ and $\boldsymbol{j}_2$, can take any value between $|j_1 - j_2|$ and $j_1 + j_2$. For two identical spins of length S the magnitude of the resultant may vary between 0 and $2S$. We shall now determine the projection operators $P_j^{(S)}(\boldsymbol{S}_1+\boldsymbol{S}_2)$ that project onto the subspace in which the eigenvalue of $(\boldsymbol{S}_1 + \boldsymbol{S}_2)^2$ is $j(j+1)$.

Starting with the relation

$$\begin{aligned} \boldsymbol{S}_1 \cdot \boldsymbol{S}_2 &= \tfrac{1}{2}\left[(\boldsymbol{S}_1 + \boldsymbol{S}_2)^2 - \boldsymbol{S}_1^2 - \boldsymbol{S}_2^2\right] \\ &= \tfrac{1}{2}\left[(\boldsymbol{S}_1 + \boldsymbol{S}_2)^2 - 2S(S+1)\right] , \end{aligned} \tag{F.3.20}$$

and exploiting the previously mentioned property that the eigenvalues of $(\boldsymbol{S}_1+\boldsymbol{S}_2)^2$ are of the form $j(j+1)$, where $j = 0, 1, 2, \ldots 2S$, the projection operator can be written in the perfectly general form

$$P_j^{(S)}(\boldsymbol{S}_1 + \boldsymbol{S}_2) = \prod_{\substack{k=0 \\ k\neq j}}^{2S} \frac{\boldsymbol{S}_1 \cdot \boldsymbol{S}_2 - x_k^{(S)}}{x_j^{(S)} - x_k^{(S)}}, \tag{F.3.21}$$

where $x_k^{(S)} = \left[k(k+1) - 2S(S+1)\right]/2$.

For two one-half spins their resultant can be a singlet state $j = 0$ and a triplet state $j = 1$. The operators that project onto the singlet and triplet subspaces are

$$\begin{aligned} P_0^{(1/2)}(\boldsymbol{S}_1 + \boldsymbol{S}_2) &= \tfrac{1}{4} - \boldsymbol{S}_1 \cdot \boldsymbol{S}_2, \\ P_1^{(1/2)}(\boldsymbol{S}_1 + \boldsymbol{S}_2) &= \tfrac{3}{4} + \boldsymbol{S}_1 \cdot \boldsymbol{S}_2. \end{aligned} \tag{F.3.22}$$

For two $S = 1$ spins the total spin can take the values $j = 0$, 1 and 2. The operators projecting on the corresponding subspaces are

$$\begin{aligned} P_0^{(1)}(\boldsymbol{S}_1 + \boldsymbol{S}_2) &= -\tfrac{1}{3} + \tfrac{1}{3}\left(\boldsymbol{S}_1 \cdot \boldsymbol{S}_2\right)^2 , \\ P_1^{(1)}(\boldsymbol{S}_1 + \boldsymbol{S}_2) &= 1 - \tfrac{1}{2}\left(\boldsymbol{S}_1 \cdot \boldsymbol{S}_2\right) - \tfrac{1}{2}\left(\boldsymbol{S}_1 \cdot \boldsymbol{S}_2\right)^2 , \\ P_2^{(1)}(\boldsymbol{S}_1 + \boldsymbol{S}_2) &= \tfrac{1}{3} + \tfrac{1}{2}\left(\boldsymbol{S}_1 \cdot \boldsymbol{S}_2\right) + \tfrac{1}{6}\left(\boldsymbol{S}_1 \cdot \boldsymbol{S}_2\right)^2 . \end{aligned} \tag{F.3.23}$$

Reference

1. A. R. Edmonds, *Angular Momentum in Quantum Mechanics*, Reissue edition, Princeton University Press, Princeton, New Jersey (1996).

Figure Credits

Fig. 2.1	J. L. Yarnell et al., *Phys. Rev. A* **7**, 2130 (1973), Figs. 3 and 4
Fig. 3.3	R. S. Preston et al., *Phys. Rev.* **128**, 2207 (1962), Fig. 4
Fig. 4.5	J. R. Chelikowsky and M. L. Cohen, *Phys. Rev. B* **14**, 556 (1976), Fig. 6
Fig. 4.7	J. R. Chelikowsky and M. L. Cohen, *Phys. Rev. B* **14,** 556 (1976), Figs. 15 and 20
Fig. 5.6	M. C. Escher's symmetry drawings. © 2006 The M. C. Escher Company-Holland. All rights reserved. www.mcescher.com
Fig. 5.27	M. C. Escher's symmetry drawings. © 2006 The M. C. Escher Company-Holland. All rights reserved. www.mcescher.com
Fig. 5.31	M. C. Escher's symmetry drawings. © 2006 The M. C. Escher Company-Holland. All rights reserved. www.mcescher.com
Fig. 7.28	R. Brill, *Solid State Physics* **20**, 1 (1967)
Fig. 8.16	B. E. Warren, *X-Ray Diffraction, Addison-Wesley, Reading, MA* (1969)
Fig. 8.17	T. M. Hayes and J. B. Boyce, *Solid State Physics* **37,** 173 (1982)
Fig. 8.18	F. Besenbacher and K. Mortensen, *Europhysics News* **21,** 68 (1990)
Fig. 10.2	S. Kugler et al., *Phys. Rev. B* **48,** 7685 (1993), Fig. 1 (upper panel) and Fig. 4 (lower panel)
Fig. 10.3	P. Lamparter and S. Steeb, *Proceedings of the Fifth Int. Conf. on Rapidly Quenched Metals, Würzburg, September 1984,* North Holland, Amsterdam (1985)
Fig. 10.4	P. Lamparter and S. Steeb, *Proceedings of the Fifth Int. Conf. on Rapidly Quenched Metals, Würzburg, September 1984,* North Holland, Amsterdam (1985)
Fig. 10.5	W. F. Nelson et al., *Phys. Rev.* **127,** 2025 (1962), Fig. 1
Fig. 10.6	D. Shechtman et al., *Phys. Rev. Lett.* **53,** 1951 (1984), Fig. 2

Fig. 10.7 M. de Boissieu et al., J. *Phys.: Condens. Matter* **2,** 2499 (1990)
Fig. 10.9 R. D. Diehl et al., J. *Phys.: Condens. Matter* **15,** R63 (2003)
Fig. 10.15 M. Senechal, *Quasicrystals and Geometry*, Cambridge Univ. Press (1995)
Fig. 10.16 P. A. Bancel et al., *Phys. Rev. Lett.* **54,** 2422 (1985), Fig. 1
Fig. 10.18 A. Katz and M. Duneau, *J. Physique* **47,** 181 (1986)
Fig. 11.11 J. W. Lynn et al., *Phys. Rev. B* **8,** 3493 (1973), Fig. 1
Fig. 11.12 G. Dolling, *Inelastic Scattering of Neutrons in Solids and Liquids*, Proceedings of the Symposium held Sept. 10–14, 1962, at Chalk River, Canada, IAEA, Vienna, 1963.
Fig. 11.13a J. L. Warren et al., *Phys. Rev.* **158,** 805 (1967), Fig. 1
Fig. 11.13b A. D. B. Woods et al., *Phys. Rev.* **131,** 1025 (1963), Fig. 1
Fig. 12.1 A. Einstein, *Annalen der Physik* **22,** 180 (1907)
Fig. 12.5 C. B. Walker, *Phys. Rev.* **103,** 547 (1956)
Fig. 12.8 P. H. Keesom and N. Pearlman, *Phys. Rev.* **91,** 1354 (1953), Fig. 1
Fig. 12.8 H. M. Rosenberg, *Low Temperature Solid State Physics*, Oxford Univ. Press (1963)
Fig. 12.9 M. Born and K. Huang, *Dynamical Theory of Crystal Lattices*, Oxford Univ. Press (1954)
Fig. 13.1 W. Krätschmer et al., *Nature* **347,** 354 (1990)
Fig. 13.3 P. Knoll and H. Kuzmany, *Phys. Rev. B* **29,** 2221 (1984), Fig. 1 (upper panel)
Fig. 13.5 T. R. Hart et al., *Phys. Rev. B* **1,** 638 (1970), Fig. 1
Fig. 13.6 P. Zhou et al., *Phys. Rev. B* **46,** 2595 (1992), Fig. 3 (lower panel)
Fig. 13.7 J. R. Hardy and A. M. Karo, *Proc. of the Int. Conf. on Light Scattering Spectra of Solids*, New York, Sept. 3–6, 1968, Springer-Verlag, Berlin (1969)
Fig. 13.9 B. P. Stoichoff et al., *Proc. of the 2nd Int. Conf. on Light Scattering in Solids,* Paris, July 19–23, 1971, Flammarion, Paris (1971)
Fig. 13.13 K.-E. Larsson et al., *Proc. of the Symp. on Inelastic Scattering of Neutrons in Solids and Liquids,* Vienna, Austria, Oct. 11–14, 1960, IAEA, Vienna (1961)
Fig. 14.5 C. G. Shull et al., *Phys. Rev.* **83,** 333 (1951), Fig. 4
Fig. 14.7 E. E. Anderson, *Proc. of the Int. Conf. on Magnetism,* Nottingham, Sept. 1964, The Institute of Physics and the Physical Society, London (1965)
Fig. 14.12 P. Weiss and R. Forrer, *Ann. de Phys.* **5,** 153 (1926)
Fig. 14.15 Y. Shapira and S. Foner, *Phys. Rev. B* **1,** 3083 (1970), Fig. 3
Fig. 15.21 M. Azuma et al., *Phys. Rev. Lett.* **73,** 3463 (1994), Figs. 2 and 3

Name Index

Abrikosov, A. A. 4
Anderson, P. W. 6, 199, 466, 585

Baker, H. F. 441
Bardeen, J. 3, 4
Becquerel, H. 52
Bednorz, J. G. 6
Berezinskii, V. L. 555
Bernal, J. D. 22
Bethe, H. 3, 564, 642
Binnig, G. 270
Blech, I. 3, 315
Bloch, F. 3, 65, 186, 509, 523, 528
Bogoliubov, N. N. 542
Born, M. 88, 185
Bose, S. N. 396
Bouckaert, L. P. 647
Bragg, W. H. 2, 241
Bragg, W. L. 2, 241, 242
Brattain, W. H. 4
Bravais, A. 113
Brillouin, L. 3, 123, 436
Brockhouse, B. N. 439
Buckminster Fuller *see* Fuller, R. Buckminster
Burgers, J. M. 285, 286

Cahn, J. W. 3, 315
Chandrasekhar, S. 24
Clausius, R. J. E. 375
Coleman, S. 201
Cooper, L. N. 4
Curie, M. 52
Curie, P. 52
Curl, R. F., Jr. 30

Darwin, C. G. 38, 259
Debye, P. 267, 389, 442
de Haas, J. W. 45
Dirac, P. A. M. 32, 37, 92, 464
Dorda, D. 6
Drude, P. 1
Dulong, P. L. 384
Dyson, F. J. 531
Dzyaloshinsky, I. E. 4

Eckart, C. 53
Einstein, A. 45, 387, 396
Escher, M. C. 115
Evjen, H. M. 84
Ewald, P. P. 84, 120, 259, 264

Faigel, G. 252
Faraday, M. 513
Fedorov, J. S. 166
Fermi, E. 71
Feynman, R. P. 532
Fibonacci 317
Frank, F. C. 290, 297
Frenkel, J. 283, 470
Friedel, G. 24
Friedrich, W. 2, 241
Fuller, R. Buckminster 29

Gerber, Ch. 270
Ghosh, D. K. 577
Ginzburg, V. L. 4

Goldschmidt, V. M. 236
Goldstone, J. 200
Gorkov, L. P. 4
Gossard, A. C. 6
Gratias, D. 3, 315
Griffith, R. B. 497
Grüneisen, E. 419, 425
Gutzwiller, M. C. 5

Haldane, F. D. M. 579
Hausdorff, F. 441
Haüy, R.-J. 109
Heisenberg, W. 3, 92, 464, 470
Heitler, W. 90
Hermann, C. 125
Higgs, P. W. 201
Holstein, T. 530
Hooke, R. 365
Hubbard, J. 5
Hund, F. 42, 96

Ising, E. 472

Jordan, P. 533
Josephson, B. D. 497

Kadanoff, L. P. 498
Kamerlingh Onnes, H. 2
Kármán, T. von 185
Kasuya, T. 466
Kerr, J. 513
Kittel, C. 465, 466
Klitzing, K. von 6
Knight, W. D. 72
Knipping, P. 2, 241
Kondo, J. 5
Korringa, J. 71
Kosterlitz, J. M. 555
Kramers, H. A. 64, 182, 466
Kronig, R. 64
Kroto, Sir H. W. 30

Lamé, G. 365
Landé, A. 54
Landau, L. D. 3, 28, 199, 489
Landsberg, G. 434
Langevin, P. 50, 53
Larmor, J. 50
Laue, M. von 2, 241
Laughlin, R. B. 6
Leggett, A. J. 4
Lehmann, O. 24
Lennard-Jones, J. E. 81, 97
Lenz, W. 472
Lieb, E. 578
Lifshitz, E. M. 200
Lindemann, F. A. 413
London, F. 78, 90
Lorentz, H. A. 2, 374, 375
Lorenz, L. V. 375
Lyddane, R. H. 377

Madelung, E. 84
Majumdar, C. K. 577
Maleev, S. V. 531
Mandelstam, F. 434
Mattis, D. C. 578
Mauguin, C.-V. 125
Mayer, J. E. 88
Mermin, N. D. 411
Miller, W. H. 119
Mössbauer, R. L. 72
Mossotti, O.-F. 375
Mott, N. F. 6
Müller, K. A. 6
Mulliken, R. S. 96

Néel, L. 453, 461, 509
Noether, E. 191

Orowan, E. 285

Pauli, W. 3
Pauling, L. C. 236
Peierls, R. E. 3, 28, 193
Penrose, R. 323
Pepper, M. 6
Petit, A. T. 384
Pisano, L. *see* Fibonacci
Planck, M. 387
Polányi, M. 285
Primakoff, H. 530
Purcell, E. M. 65

Quate, C. F. 270

Racah, G. 672
Raman, C. V. 434
Rashba, E. I. 38
Reinitzer, F. 24

Rohrer, H. 270
Ruderman, M. A. 465, 466
Rushbrook, G. S. 497
Russell, H. N. 40

Sachs, R. G. 377
Saunders, F. A. 40
Scherrer, P. 267
Schoenflies, A. M. 125, 166
Schrieffer, J. R. 4
Schrödinger, E. 32
Schultz, T. D. 578
Schwinger, J. 532
Seitz, F. 3, 117, 157
Shechtman, D. 3, 315
Shockley, W. B. 4, 295
Shubnikov, A. V. 167
Shull, C. G. 439
Sklodowska, M. *see* Curie, M.
Slater, J. C. 3, 101
Smalley, R. E. 30
Smoluchowski, R. 647
Sommerfeld, A. 3
Störmer, H. L. 6

Taylor, G. I. 285
Tegze, M. 252
Teller, E. P. 377
Thomson, J. J. 1
Thouless, D. J. 555
Tomonaga, S. 532
Tsui, D. C. 6

van der Waals, J. D. 78
Van Hove, L. 406, 656
Van Vleck, J. H. 6, 60
Voigt, W. 368
Volterra, V. 284
Voronoi, G. 117

Wagner, H. 411
Waller, I. 442
Weiss, Ch. S. 119
Weiss, P. 474
Widom, B. 498
Wigner, E. P. 3, 53, 117, 172, 533, 647
Wilson, A. H. 3
Wilson, K. G. 500

Yosida, K. 466

Zeeman, P. 2
Zener, C. 468

Subject Index

4-methoxy-benzilidene-4-butyl-aniline 24

A1 structure 204, 217
A15 structure 204, 209
A2 structure 204, 213
A3 structure 204, 226
A3′ structure 204, 226
A4 structure 204, 221
A8 structure 204, 231, 232
A9 structure 204, 234
absorption edge 269
acoustic branch 343
acoustic vibrations 340, 360–361
actinoids 179, 218
AFM *see* atomic force microscope
A_h structure 204, 207
$AlFe_3$ structure 220
alkali halides 76, 83, 219, 282
alkali metals 213
alkaline-earth metals 213, 218
almost periodic functions 311
amorphous materials 21, 303
angular frequency 339
angular momentum 665–672
anharmonicity 421
anisotropy
  ∼ constant 507
  magnetic 471–473, 504–513
annihilation operator
  of antiferromagnetic magnon 542
  of magnon 525
  of phonon 394–395
anomalous dimension 499
antibonding state 98
antiferromagnetic ground state *see* ground state, of antiferromagnet
antiferromagnetic materials 453–459
antiferromagnetic ordering
  Néel type
    absence of in one-dimensional systems 551
    presence of in two-dimensional systems 551
antiferromagnetism 453–459
  in the Heisenberg model 464
antiferromagnon *see* magnons, antiferromagnetic
antiperiodic boundary conditions *see* boundary conditions, antiperiodic
antiphase boundary 300, 301
antiphase domains 301
antisite defects 283
anti-Stokes component 434
antiunitary operator 182
atomic force microscope 270
atomic form factor 249
atomic packing factor *see* packing fraction

B1 structure 204, 219
B2 structure 204, 207, 208
B3 structure 204, 221
B4 structure 204, 228
B8 structure 204, 227, 228
Baker–Hausdorff formula 441
basis 114

BCS theory 4
benzene-hexa-n-alkanoate 24
Berezinskii–Kosterlitz–Thouless phase 556
Berezinskii–Kosterlitz–Thouless transition 555
  critical exponent in *see* critical exponent, in Berezinskii–Kosterlitz–Thouless transition
Bernal model 22, 23, 304
Bernoulli numbers 618
Bessel functions 620
Bethe ansatz 564–566
biaxial nematic phase 26
BiF_3 structure 220
bipartite lattice 520
BKT phase *see* Berezinskii–Kosterlitz–Thouless phase
black-and-white group 167
black-and-white lattice 167
Bloch equations 65
Bloch $T^{3/2}$ law 528
  corrections to 529
Bloch's theorem 186
Bloch wall 509
body-centered cubic structures 210–214
Bogoliubov transformation
  application of 546
  for antiferromagnets 542
Bohr magneton 44, 588
Bohr radius 50, 588
bonding state 98
Born–Mayer approximation 88
Born–von Kármán boundary condition *see* boundary conditions, Born–von Kármán
Bose–Einstein statistics 396
boundary conditions
  antiperiodic 186
  Born–von Kármán 185, 186, 338
  twisted 186
Bragg condition 243
Bragg peaks 243
  shape of 252
  temperature dependence of 443–444
Bragg plane 245
Bravais cell 117, 118, 144
Bravais group 140
Bravais lattices 113
  types in three dimensions 141, 146–154
  types in two dimensions 140, 142–145
Brillouin function 57
Brillouin scattering 436–437
Brillouin zone 123, 190
  of bcc lattice 212
  of fcc lattice 216
  size of, in antiferromagnets 519
buckyball 30
bulk defects *see* volume defects
bulk modulus 365, 369
bulk susceptibility *see* susceptibility, volume
Burgers circuit 286, 287
Burgers dislocation *see* screw dislocation
Burgers vector 286, 287

C1 structure 204, 219
C14 structure 204
C15 structure 204, 223
$C1_b$ structure 223
C2 structure 204
C3 structure 204, 209, 210
C4 structure 204
calamitic nematic phase 25
Cauchy relations 369
ccp structure *see* cubic crystal structures, close-packed
center-of-mass order 19
cesium chloride
  Madelung energy of 87
  structure 204, 207
chain-like structures 229–233
character table
  of O_h group 648
  of the double group $O^\dagger$ 648
chemical shift 72
chiral nematic phase *see* cholesteric phase
cholesteric phase 26
Clausius–Mossotti relation 375
Clebsch–Gordan coefficients 671
close packing
  cubic 217
  hexagonal 226

closure domains 511
coherent scattering 439–443
cohesive energy 77
  of covalent crystals 105
  of ionic crystals 84
  of molecular crystals 81–83
coincident-site lattice 299, 300
Coleman's theorem 201, 411
color center 282
color group
  black-and-white 167
  gray 167
columnar phase 28
compatibility condition 639
compressibility 365
compressional waves 366
compression modulus
    see bulk modulus
condensed phases 13
configurational entropy 276
configuration interaction 100
constant
  stiffness *see* stiffness constant
continuous groups 642–646
conventional unit cell 117, 144
coordination number 208
correlation function 16
correlation length 494–496
coupling
  jj 40
  LS 40
  Russell–Saunders 40
covalent bond 89–105
covalent crystal 89
covalent radius 233
creation operator
  of antiferromagnetic magnon 542
  of magnon 525
  of phonon 394–395
critical exponent 491, 496–500
  in Berezinskii–Kosterlitz–Thouless
    transition 555
critical temperature
  of antiferromagnets 481
  of ferrimagnets 461
  of ferromagnets 475
cross section 653–656
  of coherent scattering 439–443
crowdion 280
crystal class 162
crystal family 141
crystal field 174
crystal-field splitting 174
crystal momentum 189
crystal structure
  types of 203
crystal systems 140
  hierarchy of, in three
    dimensions 156
  hierarchy of, in two dimensions 155
crystallite 21
CsCl structure 208, 237
Cu structure 217
Cu_3Au structure 204, 208
CuAu structure 300
cubic crystal structures 205–224
  body-centered 210–214
  close-packed 217
  face-centered 214–221
  simple 205–210
cubic lattice 153
cuprite structure 204, 210
Curie's law 52
  deviation from 182
Curie susceptibility 53
Curie temperature 450, 475
  of ferrimagnetic materials 461
  of ferromagnetic materials 451–453
Curie–Weiss law 476
cyclic group 130

D0 structure 204
$D0_3$ structure 220
$D0_9$ structure 209
D2 structure 204
$D2_1$ structure 209
$D2_3$ structure 221
$D2_e$ structure 209
$D2_f$ structure 220
D8 structure 204
dangling bonds 305
Debye frequency 401
Debye function 416, 611
Debye model 389–390
Debye–Scherrer method 267
Debye temperature 412, 596
Debye–Waller factor 442

defects *see also* line defects; planar defects; point defects; volume defects
  dislocations 283–292
  grain boundaries 298
degeneracy
  accidental 173
  lifting of 173
delta function 615
dense random packing model 304
density of states
  integrated 399
  of phonons 398–409
destruction operator
  see annihilation operator
dhcp structure *see* hexagonal crystal structures, double close-packed
diamond structure 204, 221–224
diffraction
  dynamical theory of 258–260
  theory of 242–260
digamma function 619
dihedral group 130
dimerized chain 337, 345–348
dipole–dipole interaction 70, 80, 463, 661
Dirac delta function 615
direct exchange
  see exchange, Heisenberg
direct lattice 122
director 25
Dirichlet's construction 117
  in reciprocal lattice 123
disclinations 290
  screw 291
  wedge 291
discotic columnar phase 28
discotic nematic phase 25, 26
dislocation line 284
dislocations 283–292
  edge 284, 285
  mixed 285, 286
  partial
    Frank 297
    Shockley 295
  screw 285
dispersion relation
  for antiferromagnetic magnons 520, 543
  for ferromagnetic magnons 518, 524
  for phonons 395
  for spinons
    in an isotropic antiferromagnet 571
  for XY model 573
distribution function
  Bose–Einstein *see* Bose–Einstein statistics
divacancy 280
domain wall 504
  width of 511
domains 508–513
double exchange 468–469
double group 179, 641–642
double hexagonal close-packed structure *see* hexagonal crystal structures, double close-packed
Drude–Lorentz model 2
Dulong–Petit law 384
dynamical interactions 534
dynamical matrix 356
dynamical structure factor 440, 659
Dyson–Maleev transformation 531
  and the Hamiltonian of interacting magnons 534

E2 structure 204
$E2_1$ structure 209
easy axis of magnetization 472
easy plane of magnetization 472
edge dislocation 284, 285
effective magnetic moment 58
effective magneton number 58
Einstein model 387–389
elastic constants
  Lamé *see* Lamé constants
  of crystals 367–371
  Voigt *see* Voigt elastic constants
elastic waves 363–367
electron paramagnetic resonance 61
electron–phonon interaction 193
electron spin resonance 61
elementary excitations
  in magnetic systems 515
  phonons *see* phonons
entropy
  of vortices in XY model 554
EPR 61

equation of state
for crystal 418–420
equipartition theorem 384
equivalent wave vectors 190
ESR 61
Euler's constant *see* Euler–Mascheroni constant
Euler's equation 510
Euler's gamma function 619
Euler–Lagrange equation 359
Euler–Mascheroni constant 620
Evjen's method 84
Ewald construction 264–265
for Laue method 266
for powder method 268
for rotating-crystal method 267
Ewald's method 84
Madelung energy of NaCl crystal 86
Ewald sphere 264
EXAFS 269, 309
exchange 463
direct *see* exchange, Heisenberg
double *see* double exchange
Heisenberg 463–464
RKKY *see* RKKY interaction
super- *see* superexchange
exchange energy 464
exchange integral 92
exchange interaction 42
extinction length 263

F-center 282
face-centered cubic structures 214–221
factorial function 619
faithful representation of point groups 130
Faraday effect 513
Fermi contact term 71, 661
Fermi integral 612
Fermi pseudopotential 247, 439
ferrimagnetic materials 461–462
ferrimagnetism 461, 462
ferromagnetic materials 450–453
ferromagnetism 450–453, 470
Fibonacci chain 317–323
fine structure 40
of paramagnetic resonance 68
fine-structure constant 36
first theorem of condensed-matter physics 199
fixed point 502
flip-over process *see* umklapp process
Floquet's theorem 186
fluctuation–dissipation theorem 52
fluorite structure 204, 219
flux quantum 587
Fourier transform 601
Frank partial dislocation 297
free energy
of vortices in XY model 555
free enthalpy *see* Gibbs free energy
Frenkel defect 283
Fresnel's equations 259
fullerene 30, 76
fullerite 30, 221
infrared absorption in 432
Raman scattering in 435

g-factor 45
gamma function 619
γ-selenium structure 204
gap
in magnon spectrum for antiferromagnets 546
Gaussian function
Fourier transform of 609
Gaussian system of units 590
Gell-Mann matrices 646
Gibbs free energy 277
Gibbs potential *see* Gibbs free energy
glass 21
metallic *see* metallic glass
spin *see* spin glass
glide line 158
glide plane 158
glide reflection 158
Goldstone bosons 200, 397
Goldstone's theorem 200, 397, 546, 547
grain boundaries 293, 298
tilt 298
twist 298
graphite structure 204
gray groups 168
Griffiths inequality 497
ground state
of antiferromagnet 543–544
and the Néel state 543–544, 568

of crystal lattice 410–413
of ferromagnet 522
group theory 633–651
Grüneisen parameter 419
Grüneisen relation 425
gyromagnetic ratio 45

H1 structure 204
$H1_1$ structure 223
Haldane gap 579
half-Heusler structure 223
Hall effect
quantum 6
Hamiltonian
of interacting magnons 534
Hankel functions 620
harmonic approximation 331–337
harmonic oscillator 392–393
Hartree energy 588
Hartree–Fock approximation 96
hcp structure *see* hexagonal crystal structures, close-packed
Heaviside step function 608, 616
HEED 262
Heisenberg chain
anisotropic ferromagnetic
spin-1/2 560–566
antiferromagnetic
spin-1 578–580
spin-1/2 566–569
Heisenberg exchange *see* exchange, Heisenberg
Heisenberg model 469
Hamiltonian
eigenstates of 521–522, 526–527, 538, 540
quadratic and quartic parts 534–535
Heitler–London approximation 90
Helmholtz free energy 278
of phonon gas 414
Hermann–Mauguin symbols 125
Hermite polynomials 623
heteropolar bond 83
Heusler alloy *see* Heusler phase
Heusler phase 204, 220
hexa-*n*-alkoxy triphenylene 24
hexagonal crystal structures 224–229
close-packed 204, 224, 226
double close-packed 224, 226
simple 224
hexatic phase 28
Higgs bosons 201
high-temperature expansion 503–504
HMTTF-TCNQ 232
Holstein–Primakoff transformation 530
and the Hamiltonian of interacting magnons 534
application of 534, 546
homeopolar bond 89
homopolar bond 89
honeycomb lattice 113, 114
Hooke's law 365, 368
Hubbard model 5
Hund's rules 42
hybrid states 103–105, 233
hydrogen bond 106
hyperfine structure 69

icosahedral group 131
ideal crystal 14, 109
improper rotation 126
incoherent scattering 660
incommensurate structures 312
magnetic *see* spiral structures
indirect exchange *see* RKKY interaction
inelastic neutron scattering *see* neutron scattering, inelastic, 547
infrared absorption 431–433
infrared active mode 432
integrated density of states 399
interaction
electron–phonon 193
phonon–phonon 423–424
interfacial defects 274
international notation *see* Hermann–Mauguin symbols
interstitials 278–280
split 279
inversion 126
ionic bond 83
ionic–covalent bond 94
ionic crystals 83–89
optical vibrations in 373–377
ionic radius 236
irreducible representations 637

Ising model 472
and Mermin–Wagner theorem 551
two-dimensional 551
isomer shift 73

jj coupling 40
Jones symbol 130
Jordan–Wigner transformation 533
application of 572
Josephson constant 587
Josephson inequality 497, 499
junction transistor 4

kagome lattice 113, 114
Kerr effect 513
kinematical interaction 536
Knight shift 72
Kondo effect 5
Korringa law 71
Korringa relaxation 71
Kosterlitz–Thouless transition *see* Berezinskii–Kosterlitz–Thouless transition
Kramers–Kronig relation 64
Kramers' theorem 182

L1 structure 204, 207
L2 structure 204
L′3 structure 204
ladder operator 393
Lagrange's equation 34, 359
Laguerre polynomials 624
Lamé constants 365
Landé g-factor 54
Landau–Peierls instability 28, 201
Landau theory of phase transitions 199, 489–492
Langevin diamagnetism *see* Larmor diamagnetism
Langevin function 57
Langevin susceptibility 53
lanthanoids 179, 218
Larmor diamagnetism 49
Larmor frequency 50
lattice parameters 229
lattice vibrations
classical description of 331–385
Einstein model 387–389
quantum description of 387–427
lattices that are not Bravais lattices 113, 114
Laue condition 194, 244, 245
Laue method 265
Laves phase 204, 223
layered structures 229–233
LCAO method 97
LEED 262
Legendre polynomials 625
Lennard-Jones potential 81
level splitting in crystals 173–182
Lie groups 643
Lieb–Schultz–Mattis theorem 578, 582
lifetime
of magnons 536, 548
of phonons 424
Lindemann criterion 413
line defects 274, 283–292
linear chain
diatomic 341–345
dimerized 345–348
monatomic 337–341
liquid crystals 24–29
liquid phase 22–23
localized excitations
lattice vibrations 377–383
long-range order 14
longitudinal vibrations 341
Lorentz formula 374
Lorentz–Lorenz equation 375
Lorentzian function
Fourier transform of 609
low-angle grain boundary 298
low-dimensional magnetic systems 548f
LS coupling 40
Lyddane–Sachs–Teller relation 377

M-center 282
Madelung constant 87
Madelung energy 84
magnetic force microscope 513
magnetic form factor 249
magnetic group *see* color group, black-and-white
magnetic lattice *see* black-and-white lattice
magnetic space groups *see* space groups, black and white

magnetic structures 453–462
magnetism
  antiferro- 453–459
  atomic 51
  ferri- 461, 462
  ferro- 450–453, 470
magnetization *see also* sublattice magnetization
  definition of 48
  temperature dependence of *see* Bloch $T^{3/2}$ law
magnetomechanical ratio *see* gyromagnetic ratio
magnon energy
  temperature-dependent corrections to 535
magnons
  antiferromagnetic 540–547
  approximate bosonic character of 521, 525
  as magnetic counterparts of phonons 521
  bound states of 536–540, 565–566
  cutoff for 527, 545
  ferrimagnetic 546–547
  ferromagnetic 521
  interaction of 533–536
    Hamiltonian for 534
  thermodynamics of 527–530
Majumdar–Ghosh point 577
mass susceptibility *see* susceptibility, mass
MBBA 24
mean free path
  of phonons 427
mean-field theory
  of antiferromagnetism 478–484
  of ferrimagnetism 487–488
  of ferromagnetism 474–478
melting point of elements 596
Mermin–Wagner theorem 411, 550
mesogen 24
mesomorphic phases 13, 23–24
mesoscopic systems 7
metallic bond 106
metallic glass 21, 303
Mg structure *see* hexagonal crystal structures, close-packed
Miller indices 119
Miller–Bravais indices 119
mirror line 124
mirror plane 124
misorientation angle 298
mixed dislocation 285, 286
$MoAl_{12}$ structure 214
modulus of elasticity *see* Young's modulus
modulus of rigidity *see* shear modulus
molar susceptibility *see* susceptibility, molar
molecular crystals 78–83
molecular-orbital method 96
monatomic chain *see* linear chain, monatomic
mosaic structure 298
Mössbauer effect 72
motif 114

N-process *see* normal process
NaCl structure *see* sodium chloride, structure
nanostructures 7
Néel state 540
Néel temperature 454, 481
Néel wall 509
nematic phase 25, 26
  biaxial 26
  calamitic 25
  chiral 26
  cholesteric 26
  discotic 25, 26
Neumann functions 620
Neumann's principle 171
neutron scattering
  cross section of *see* Van Hove formula
  elastic 241
  inelastic 242, 438–447
  magnetic 660–664
NiAs structure 227, 228
NMR *see* nuclear magnetic resonance
noble gases 77, 218
noble metals 218
Noether's theorem 191
noncrystalline solids 21–22
nonsymmorphic plane groups 163
nonsymmorphic space groups 166
normal coordinates 357

normal modes 359
normal process 193
nuclear magnetic resonance 71

octahedral group 130
octahedral sites 213, 218
optical branch 343
optical vibrations 343, 361–363
  in ionic crystals 373–377
order
  center-of-mass 19
  long-range 14
    absence of in finite temperature XY model 555
    absence of in low-dimensional finite temperature Heisenberg model 550
  magnetic 450–462
  orientational 19
  short-range 14
order–disorder transitions 489
orientational order 19
oscillator 392–393
overlap integral 92

PAA 24
packing fraction
  definition of 208
  for ccp crystal 217
  for diamond structure 222
  for hcp crystal 226
  for sc crystal 208
pair-correlation function 17
para-azoxy anisole 24
parallel susceptibility 484
paramagnetic resonance 61
  fine structure 68
paramagnetic space groups *see* space groups, gray
paramagnetism 51
  atomic 51
  Van Vleck 60
partial dislocations 294–298
partial structure factor 307
particle–hole excitations
  in the XY model 574–575
Pauli exclusion principle 36, 464
Pauli matrices 646, 674
Pauling ionic radius 236
Pearson symbol 205
Penrose tiling 323–326
periodic boundary condition *see* boundary conditions, Born–von Kármán
periodic table 593
perovskite structure 204, 209
perpendicular susceptibility 484
phonon–phonon interaction 423–424
phonons 395
  acoustic 397
  density of states of 398–409
  experimental study of 429–447
  interaction among *see* phonon–phonon interaction
  lifetime of 446
  specific heat of 413–418
physical constants 587–588
planar defects 274, 293–301
planar model *see* XY model
planar regime 566
plane groups 162
plastic crystals 29–30
point defects 274–283
point group of the crystal 161
point groups 127–135
point-contact transistor 4
Poisson's ratio 369
polar covalent bond 94
polarization vector 357
polycrystals 21, 293
polymers 30
powder method 267
primitive cell 115
primitive vectors 110
  choice of 111–113
pyrite structure 204

quantum critical point 500
quantum Hall effect 6
quantum phase transitions 500
quasicrystals 3, 21, 315–330
quasimomentum 191
quasiperiodic functions 311
quasiperiodic structures 309–330
quasiperiodic tiling *see* Penrose tiling

R-center 282
radial distribution function 17, 305

in amorphous silicon 306
in quasicrystals 316
Raman active mode 435
Raman scattering 433–436
rapidity 563
rare-earth garnet 547
rare-earth metals 452
Rashba term 38
reciprocal lattice 120–124
definition 120
of bcc lattice 212
of fcc lattice 216
of hexagonal lattice 225
primitive vectors of 122
reduction 638
relativistic effects 36–38
relaxation function 63
relaxation time
spin–lattice 64
spin–spin 65
renormalization 501
renormalization-group transformation 500–502
resonance absorption 61, 72
resonance fluorescence 72
resonance integral 98
resonating valence bond spin liquid 585
Riemann zeta function 617
RKKY interaction 466
rock-salt structure *see* sodium chloride, structure
Rodrigues' formula
for generalized Laguerre polynomials 625
for Hermite polynomials 623
for Laguerre polynomials 624
rotating-crystal method 265
rotation axis 125
rotation group 642, 665
rotation–inversion *see* symmetry operations, rotation–inversion
rotation–reflection *see* symmetry operations, rotation–reflection
rotational symmetry 125
rotoinversion *see* symmetry operations, rotation–inversion
rotoreflection *see* symmetry operations, rotation–reflection
Ruderman–Kittel oscillation 465
Rushbrook inequality 497
Russell–Saunders coupling 40
rutile structure 204
RVB *see* spin liquid, resonating valence bond
Rydberg energy 588

satellite peaks 314, 459
saturated bond 99
scaling laws 496–500
scanning tunneling microscope 270
scattering length 247, 440
scattering vector 548
Schoenflies symbols 125
Schottky defect 280–282
Schwinger boson 532
screw axis 158
screw dislocation 285
screw rotation 158
s–d interaction 465
Seitz symbol 157
selection rules 184, 650
separation energy 77
shear modulus 365, 369
shift operator *see* ladder operator
Shockley partial dislocation 295
short-range order 14, 305–309
Shubnikov groups 167
simple cubic crystals 205–210
single crystal 20
skutterudite structure 204
Slater determinant 101
smectic phases 27–28
sodium chloride
Madelung energy of 86
structure 204, 219, 238
solid solutions 21
space groups 162
black-and-white 168
gray 168
in two dimensions *see* plane groups
magnetic 166
specific heat
of classical crystals 383–385
of magnon gas 529
of phonon gas 413–418
specific susceptibility *see* susceptibility, specific

spectrum
in chain of $S = 1/2$ spins 539
in the XY model 575
sphalerite structure 204, 221–224
spherical Bessel functions 621
spherical Hankel functions 622
spherical harmonics 627
spherical Neumann functions 622
spin
classical
precession of 516
classical equation of motion in magnetic field 516
quantum mechanical equation of motion 516
spin-flop phase 484
spin-flop transition 484, 546
spin glass 449
spin ladders 581–583
spin–lattice relaxation 64
spin liquid 551, 574 584–586
algebraic 584
resonating valence bond 585
spin operators 665
representation in terms of boson operators 530–533, 541
representation in terms of fermion operators 533
spin–orbit interaction 38
spin waves
antiferromagnetic 540–547
classical 516–520
antiferromagnetic 518–520
ferromagnetic 516–518
illustration of 518, 520
definition of 517
ferromagnetic 521
quantum mechanical description of 521–540
spinel structure 204, 223
spinons 570
dispersion relation of *see* dispersion relation, of spinons
spiral structures 459–461
split interstitial 279
stacking faults 293–294
staggered susceptibility 482
step function *see* Heaviside step function
stiffness constant
of spin waves 524
Stirling formula 619
STM *see* scanning tunneling microscope
Stokes component 434
strain tensor 364
stress tensor 365
structure amplitude 248
structure factor 18, 248
dynamical *see* dynamical structure factor
in amorphous silicon 306
partial 307
Strukturbericht designation 205
sublattice magnetization *see also* magnetization
temperature dependence of 545
substitutional impurity 274
superconductivity 2
superexchange 466–468
susceptibility
Curie 53
definition of 48
Langevin 53
mass 49
molar 49
of antiferromagnets 482–484
of ferromagnets *see* Curie–Weiss law
divergent spin-wave contribution to 530
specific 49
Van Vleck 60
volume 49
symmetry breaking 199
symmetry elements
glide line 158
glide plane 158
inversion center 126
mirror line 124
mirror plane 124, 125
rotation axis 125
rotation–inversion axis 126
rotation–reflection axis 126
screw axis 158
symmetry operations
glide reflection 158
inversion 126

reflection 125
rotation–inversion 126
rotation–reflection 126
screw rotation 158
symmorphic plane groups 163
symmorphic space groups 164
synchrotron radiation 241
syngony *see* crystal systems

Taylor–Orowan dislocation *see* edge dislocation
TBBA 24
tensile modulus *see* Young's modulus
terephtal-bis-butyl-aniline 24
tetrahedral group 130
tetrahedral sites 213, 218
thermal conductivity
by phonons 427
thermal expansion 425–427
tilt grain boundary 298
time reversal 196
top-hat function
Fourier transform of 608
topological quantum number 554
torsional waves 366
transistor
junction 4
point-contact 4
transition metals 213, 218
translational symmetry 110
transverse vibrations 341
triple-axis spectrometer 439, 445
TTF-TCNQ 76
twin crystals 299, 301
twist grain boundary 298
twisted boundary conditions *see* boundary conditions, antiperiodic
two-phonon Raman scattering 436

umklapp process 193, 424, 427
uniaxial anisotropy 472
unsaturated bond 99

vacancies 275–278
formation energies of 276
vacancy pair 280
valence-bond method 90
valence-bond-solid state 580
van der Waals bond 79–81
van der Waals interaction 78
Van Hove formula 440, 656
Van Hove singularities 405–409
Van Vleck paramagnetism 60
Van Vleck susceptibility 60
vector operator
definition of 53
Voigt elastic constants 368
Volterra construction 284
volume defects 274, 302
volume susceptibility *see* susceptibility, volume
von Klitzing constant 587
Voronoi polyhedron 117
vortices 553–559

W structure 213
wallpaper groups *see* plane groups
Weiss field 474
Weiss indices 119
Wigner–Eckart theorem 53
Wigner–Seitz cell 117
Wigner–Seitz sphere 116
Wigner's theorem 173
wurtzite structure 204, 228

X-ray diffraction
experimental methods of 261–268
theory of 242–260
XY model 472, 551, 572
entropy of vortices in *see* entropy, of vortices in XY model
free energy of vortices in *see* free energy, of vortices in XY model
Hamiltonian of 551
phase transition in *see* Berezinskii–Kosterlitz–Thouless transition

YIG *see* yttrium–iron garnet
Young's modulus 369
yttrium–iron garnet 547
Yukawa function
Fourier transform of 609

zero-point energy 392, 543
zero-point spin contraction
in antiferromagnetic ground state 544
in two-dimensional antiferromagnets 551
zero-point vibrations 89, 392
zeta function 617
zincblende structure *see* sphalerite structure

Fundamental physical constants

Name	Symbol	Value
Bohr magneton	$\mu_{\mathrm{B}} = e\hbar/2m_{\mathrm{e}}$	$9.274\,009 \times 10^{-24}\ \mathrm{J\,T^{-1}}$
Bohr radius	$a_0 = 4\pi\epsilon_0\hbar^2/m_{\mathrm{e}}e^2$	$0.529\,177 \times 10^{-10}\ \mathrm{m}$
Boltzmann constant	k_{B}	$1.380\,650 \times 10^{-23}\ \mathrm{J\,K^{-1}}$
Conductance quantum	$G_0 = 2e^2/h$	$7.748\,092 \times 10^{-5}\ \mathrm{S}$
Electron g-factor	$g_{\mathrm{e}} = 2\mu_{\mathrm{e}}/\mu_{\mathrm{B}}$	$-2.002\,319$
Electron gyromagnetic ratio	$\gamma_{\mathrm{e}} = 2\lvert\mu_{\mathrm{e}}\rvert/\hbar$	$1.760\,860 \times 10^{11}\ \mathrm{s^{-1}\,T^{-1}}$
	$\gamma_{\mathrm{e}}/2\pi$	$28\,024.9540\ \mathrm{MHz\,T^{-1}}$
Electron magnetic moment	μ_{e}	$-9.284\,764 \times 10^{-24}\ \mathrm{J\,T^{-1}}$
		$-1.001\,160\,\mu_{\mathrm{B}}$
Electron mass	m_{e}	$9.109\,382 \times 10^{-31}\ \mathrm{kg}$
Electric constant	$\epsilon_0 = 1/\mu_0 c^2$	$8.854\,188 \times 10^{-12}\ \mathrm{F\,m^{-1}}$
Elementary charge	e	$1.602\,176 \times 10^{-19}\ \mathrm{C}$
Hartree energy	$E_{\mathrm{h}} = e^2/4\pi\epsilon_0 a_0$	$4.359\,744 \times 10^{-18}\ \mathrm{J}$
in eV		$27.211\,383\ \mathrm{eV}$
Josephson constant	$K_{\mathrm{J}} = 2e/h$	$483\,597.9 \times 10^{9}\ \mathrm{Hz\,V^{-1}}$
Magnetic constant	μ_0	$4\pi \times 10^{-7}\ \mathrm{N\,A^{-2}}$
Magnetic flux quantum	$\Phi_0 = h/2e$	$2.067\,834 \times 10^{-15}\ \mathrm{Wb}$
Nuclear magneton	$\mu_{\mathrm{N}} = e\hbar/2m_{\mathrm{p}}$	$5.050\,783 \times 10^{-27}\ \mathrm{J\,T^{-1}}$
Neutron mass	m_{n}	$1.674\,927 \times 10^{-27}\ \mathrm{kg}$
Neutron magnetic moment	μ_{n}	$-0.966\,236 \times 10^{-26}\ \mathrm{J\,T^{-1}}$
		$-1.913\,043\,\mu_{\mathrm{N}}$
Neutron g-factor	$g_{\mathrm{n}} = 2\mu_{\mathrm{n}}/\mu_{\mathrm{N}}$	$-3.826\,085$
Planck constant	h	$6.626\,069 \times 10^{-34}\ \mathrm{J\,s}$
in eV	$h/\{e\}$	$4.135\,667 \times 10^{-15}\ \mathrm{eV\,s}$
Proton g-factor	$g_{\mathrm{p}} = 2\mu_{\mathrm{p}}/\mu_{\mathrm{N}}$	$5.585\,695$
Proton gyromagnetic ratio	$\gamma_{\mathrm{p}} = 2\mu_{\mathrm{p}}/\hbar$	$2.675\,222 \times 10^{8}\ \mathrm{s^{-1}\,T^{-1}}$
	$\gamma_{\mathrm{p}}/2\pi$	$42.577\,482\ \mathrm{MHz\,T^{-1}}$
Proton magnetic moment	μ_{p}	$1.410\,607 \times 10^{-26}\ \mathrm{J\,T^{-1}}$
		$2.792\,847\,\mu_{\mathrm{N}}$
Proton mass	m_{p}	$1.672\,622 \times 10^{-27}\ \mathrm{kg}$
Reduced Planck constant	$\hbar = h/2\pi$	$1.054\,572 \times 10^{-34}\ \mathrm{J\,s}$
in eV	$\hbar/\{e\}$	$6.582\,119 \times 10^{-16}\ \mathrm{eV\,s}$
Rydberg constant	$R_\infty = \alpha^2 m_{\mathrm{e}}c/2h$	$10\,973\,731.569\ \mathrm{m^{-1}}$
Rydberg energy	$\mathrm{Ry} = R_\infty hc$	$2.179\,872 \times 10^{-18}\ \mathrm{J}$
in eV		$13.605\,692\ \mathrm{eV}$
Speed of light	c	$299\,792\,458\ \mathrm{m\,s^{-1}}$
Von Klitzing constant	$R_{\mathrm{K}} = h/e^2$	$25\,812.807\,572\ \Omega$